



HAL
open science

Un système d'information géographique pour le suivi d'objets historiques urbains à travers l'espace et le temps

Bertrand Dumenieu

► To cite this version:

Bertrand Dumenieu. Un système d'information géographique pour le suivi d'objets historiques urbains à travers l'espace et le temps. Histoire. École des Hautes Études en Sciences Sociales, 2015. Français. NNT: . tel-03262391

HAL Id: tel-03262391

<https://theses.hal.science/tel-03262391>

Submitted on 16 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Thèse

pour obtenir le grade de docteur de l'EHESS

Mention : Histoire

Spécialité : Mathématiques et applications aux sciences de l'homme

Un système d'information géographique pour le suivi d'objets historiques urbains à travers l'espace et le temps

-

Bertrand Dumenieu



Soutenue publiquement le 4 décembre 2015

Composition du jury :

Christophe Claramunt	Professeur des Universités, École Navale	Rapporteur
Sophie de Ruffray	Professeur des Universités, Université de Rouen	Rapporteur
Lorenz Hurni	Professeur, ETH Zürich	Examineur
Jean-Pierre Nadal	Directeur de Recherches, CNRS & Directeur d'Études, EHESS	Examineur
Cyril de Runz	Maître de Conférences (HDR), IUT de Reims-Châlons-Charleville	Invité
Maurizio Gribaudo	Directeur d'Études, EHESS	Directeur de thèse
Anne Ruas	ICPEF (HDR), IFSTTAR	Directrice de thèse
Julien Perret	Chargé de Recherches, IGN	Encadrant

Table des matières

Introduction	5
I Positionnement et sources cartographiques sur Paris	9
1 Positionnement du sujet	11
1.1 Analyser l'espace pour comprendre l'évolution d'une société urbaine	14
1.2 Représenter un espace urbain en transformation	23
1.3 L'apport des SIG pour étudier l'histoire : un choix méthodologique	38
1.4 Les difficultés dues aux imperfections de l'information géohistorique	44
1.5 État de l'art des SIG utilisés en histoire	50
1.6 Positionnement	57
1.7 Conclusion	61
2 Sources cartographiques sur Paris	63
2.1 Atlas de Verniquet	67
2.2 Plan de Maire	69
2.3 Cadastre de Vasserot	73
2.4 Atlas de Jacoubet	75
2.5 Atlas municipal	77
2.6 L'atlas de Jacoubet : des architectes voyers à l'administration de la Seine	80
2.7 Conclusion	107
II Construction de graphes spatio-temporels	109
3 Géoréférencement des plans de Paris	111
3.1 Construire une carte topographique au XIX ^e siècle	114
3.2 Géoréférencement : principe de la méthode usuelle	120
3.3 Évaluation de la qualité d'un géoréférencement indirect	129
3.4 Géoréférencement des plans de Paris	133
3.5 Cas applicatif : l'atlas de Verniquet	147
3.6 Vectorisation des plans	171
3.7 Conclusion	174
4 Modèle de données spatial et temporel	175
4.1 Modélisation de la base de données spatiale et temporelle	177
4.2 Représentation du temps dans les bases de données temporelles	188

4.3	Représentation floue du temps	189
4.4	Tri temporel	198
4.5	Conclusion	208
5	Un graphe pour modéliser les transformations de l'espace	209
5.1	Modèle de graphe spatio-temporel de Del Mondo	211
5.2	Adaptation du graphe spatio-temporel aux observations géohistoriques . . .	221
5.3	Processus spatio-temporels	230
5.4	Conclusion	236
6	Construction du graphe géohistorique	237
6.1	Approche générale	240
6.2	Découverte des relations de filiation	248
6.3	Typage des relations de filiations	277
6.4	Mise en œuvre et résultats : application aux rues de Paris	281
6.5	Conclusion	306
	Conclusion et perspectives	309
	Annexes	317
A	Carte des points de triangulation de l'atlas de Verniquet	319
B	<i>Atlas National de la Ville de Paris</i>	323
C	<i>La topographie de Paris</i>	325
D	<i>Atlas Général de la Ville et des Faubourgs de Paris</i>	327
E	<i>Atlas municipal des vingt arrondissements de la ville de Paris</i>	329
F	Agrandissements des réseaux de rue vectorisés	331
G	Indices temporels des planches de l'Atlas de Jacoubet	337
H	Évaluation des méthodes de rangement de quantités floues	341
I	Mesures de l'antécédence de deux temps flous	345
	Bibliographie	347
	Table des Figures	375
	Liste des Tableaux	377

Introduction

Depuis le milieu des années 1990, les disciplines historiques montrent un intérêt grandissant pour les systèmes d'informations géographiques (SIG). En effet, ceux-ci offrent la possibilité de produire des représentations visuelles des espaces passés, soit à partir de représentations cartographiques anciennes, soit en constituant des bases de données spatiales à partir de sources diverses (cartes, données archéologiques, connaissances historiques, etc.). Les données produites par des études historiques sur des individus, des événements, ou encore des zones du territoire considérés à un temps donné peuvent alors être associées à des références spatiales issues de ces bases de données représentant l'espace passé à une période cohérente avec le temps valide des données. L'inscription de ces données historiques dans l'espace fournit ainsi de nouvelles perspectives d'analyse et de compréhension de la répartition et de l'organisation des phénomènes étudiés ainsi que de leurs interactions avec l'espace. Enfin, les SIG constituent de puissants outils de partage des données utiles aux études historiques et de diffusion des connaissances produites par ces études. Les plate-formes cartographiques en ligne et les bases de données spatiales ouvertes offrent de nouveaux outils de diffusion, de visualisation et d'analyse des données historiques spatialisées. Elles augmentent en outre la ré-utilisabilité de ces données. Les phénomènes étudiés dans le champs de l'histoire sont par essence dynamiques, et leur étude se situe dans la durée. Spatialisés, ils évoluent dans un espace lui-même en constante réorganisation. Ainsi, une représentation statique ne suffit plus et il devient nécessaire de représenter également les changements morphologiques qui affectent cet espace au cours du temps. Lorsque les phénomènes dont l'évolution au cours du temps est étudiée sont purement spatiaux, comme les transformations d'une ville par exemple, la représentation des dynamiques spatiales au sein d'un système d'information géographique offre des outils de visualisation et de mesure de ces changements. Lorsque qu'ils sont sociaux, la confrontation entre dynamiques spatiales et sociales permet de faire apparaître les interactions qui peuvent exister entre elles (Knowles et Hillier, 2008).

Du point de vue de la géomatique, cette question de la représentation des évolutions de l'espace se traduit par la production de bases de données spatio-temporelles. Celles-ci permettent de représenter des objets géographiques évoluant dans le temps et fournissent un ensemble d'outils permettant d'effectuer des analyses sur ces évolutions.

Pour former une base de données géographique sur l'espace ancien à une période choisie, deux stratégies sont possibles. La première consiste à partir d'une base de données de référence, puis à modifier les objets qui la composent selon les connaissances historiques dont on dispose sur leur état et leur existence au moment considéré. La seconde stratégie repose sur des sources cartographiques anciennes qui sont numérisées de façon à extraire des objets géographiques. La première stratégie n'est envisageable que pour des périodes

proches de celle de la base de données de référence pour lesquelles les structures sont connues et peu différentes. Pour cette raison, nous adoptons dans cette thèse uniquement sur la seconde stratégie. Celle-ci comporte cependant deux inconvénients majeurs. Tout d'abord, les sources cartographiques anciennes sont hétérogènes : leur forme, leur objectif et leur contexte différent. En résultent des choix d'échelles, de thèmes cartographiques ou de symbolisation variés. Même lorsqu'elles décrivent un même espace, les pratiques cartographiques contemporaines de leur production aboutissent à des divergences majeures entre cartes comme le recours à des vues obliques ou verticales par exemple. De plus, des différences de précision géométrique ainsi que de perfectionnement des techniques de levé cartographique aboutissent à des cartes distordues pour lesquelles une simple superposition ne suffit pas à saisir les transformations de l'espace tracé. Les objets géographiques numérisés à partir de sources cartographiques anciennes sont donc, à leur image, hétérogènes.

La constitution d'une base de données spatio-temporelle sur l'espace nécessite de situer les objets géographiques dans le temps, cette temporalité dépendant de la nature des objets et de l'étude. Cette datation, effectuée à partir de connaissances historiques éventuellement partielles ou incertaines, peut présenter des difficultés. En outre, lorsque les objets sont issus de sources cartographiques, il faut se poser la question de leur statut. S'agit-il d'objets géographiques représentant des entités du monde réel et dont la nature influe sur la datation ? Ou bien doit-on les considérer comme de simples symboles dont l'existence et donc la date sont intrinsèquement liées à celles de la carte qui les renferme ? Il faut alors s'interroger sur la temporalité propre à la source cartographique elle-même : quelle signification donner à l'existence d'une carte ?

Il existe de nombreux systèmes d'information géographique dédiés à l'histoire. Pour la plupart, il s'agit d'outils de cartographie en ligne proposant une visualisation de données sociales localisées sur un fond cartographique ancien. Seuls quelques outils modélisent les transformations et s'appuient sur une base de données spatio-temporelle. Ils sont cependant spécialisés pour un cas applicatif précis. L'absence d'outils génériques pour créer de telles bases de données s'explique par les différentes difficultés que nous venons d'énoncer. L'inadéquation entre les fonctionnalités des SIG existants et les spécificités des données historiques constituent un frein à la mise en œuvre de tels outils. L'objectif de cette thèse est donc de **proposer une méthode de création de bases de données spatio-temporelles à partir de sources cartographiques anciennes.**

Étant donnée la position interdisciplinaire du sujet qu'il aborde, ce mémoire s'organise en deux parties, chacune adressée à un des champs disciplinaires concernés.

La première partie divisée en deux chapitres, s'inscrit dans le domaine de l'histoire. Dans le **chapitre 1**, nous posons les bases méthodologiques d'un modèle de création de données spatio-temporelles pour l'étude historique. Pour ce faire, nous nous intéressons dans un premier temps aux questionnements des historiens sur l'espace et ses transformations afin de dégager les deux objectifs généraux d'un outil destiné à modéliser ces transformations. En analysant un ensemble de travaux historiques, nous identifions trois niveaux d'intrication entre phénomènes sociaux et dynamiques spatiales mettant en évidence le besoin de données spatio-temporelles

pour décrire ces dynamiques, mais également d'outils permettant de les manipuler. Nous proposons ensuite une typologie des sources d'archives permettant de retrouver les transformations de l'espace, avant de nous concentrer plus particulièrement sur les plans topographiques. Nous montrons par la suite les différentes complexités et imperfections spatiales mais aussi temporelles de ces sources. De tous ces éléments, nous dégagons les trois objectifs d'un outil SIG adapté à l'histoire :

- La constitution de données spatio-temporelles pour représenter les transformations de l'espace,
- La modélisation et la représentation des imperfections des informations sur l'espace,
- La conservation d'un lien fort entre les données et les sources historiques d'archives desquelles elles proviennent.

Nous confrontons ensuite les différents SIG historiques avec les objectifs identifiés afin de justifier notre approche méthodologique ainsi que les objectifs de cette thèse.

Dans le **chapitre 2**, nous proposons une analyse des sources cartographiques parisiennes utilisées pour extraire des objets géographiques sur l'espace de Paris afin d'identifier leurs imperfections propres mais aussi leurs spécificités et leur contenu. Ce chapitre est également l'occasion d'effectuer une critique de ces différentes sources, en particulier dans le cas de l'*atlas national de la ville de Paris* dressé par l'architecte Théodore Jacoubet. L'étude de cet atlas nous permet en outre de mettre en lumière l'organisation progressive des services de la voirie parisienne au cours du XIX^e siècle. Enfin, ce chapitre nous permet de mieux connaître les temporalités de ces différentes sources cartographiques.

La seconde partie présente le cœur de notre proposition au fil de 4 chapitres. Ceux-ci répondent à la question générale de la construction d'une base de données spatio-temporelle à partir de représentation de l'espace urbain décrit par des sources cartographiques anciennes?. Nous décomposons la question en 4 sous-questions, chacune traitée par un chapitre.

Dans le **chapitre 3**, nous abordons la problématique du géoréférencement de plans topographiques anciennes de Paris, ainsi que de l'extraction de données géographiques vectorielles qui puissent être utilisées pour reconstituer les transformations de la ville. Pour cela, nous proposons une approche de géoréférencement fondée sur l'utilisation des éléments géodésiques représentés dans les plans, notamment le carroyage de ceux-ci. Cette approche nous permet, à partir de points de mesure placés sur des objets géographiques, d'analyser en finesse les distorsions géométriques de l'atlas de Verniquet, un grand plan de Paris du XVIII^e.

Le **chapitre 4** présente quand à lui le modèle de données spatiales et temporelles que nous avons créé, destiné à stocker les données géographiques vecteur issues des cartes. Ce chapitre s'organise autour de deux axes. D'abord, nous proposons un modèle de base de données fondée sur les standards de l'information géographique visant à représenter des objets géographiques vectoriels extraits de sources cartographiques tout en conservant une relation de dépendance forte entre ces objets et leur source. En outre, nous présentons comment ces objets, nommés **observations géohistoriques** dans le modèle, peuvent être localisés dans le temps. Le modèle

finalement proposé organise les observations en couches spatiales et temporelles, ou *snapshots*. La localisation temporelle des observations et de leurs sources étant souvent difficile à déterminer avec précision, nous proposons une représentation des temporalités par intervalles temporels flous décrivant la période d'existence des sources et des observations géohistoriques. Nous décrivons également la méthodologie adoptée pour définir ces intervalles flous, ce qui nous permet de fixer les temporalités associées aux sources cartographiques utilisées. Enfin, à partir d'un état de l'art des méthodes de tri de sous-ensembles flous, nous sélectionnons deux opérateurs de base pour raisonner sur des temporalités floues, en particulier sur l'ordre temporel et l'antécédence d'un temps sur un autre.

Dans le **chapitre 5**, nous nous intéressons à la question suivante : *comment modéliser et identifier les transformations existantes entre les observations de différentes sources géohistoriques ?*. Nous y répondons en deux temps. D'abord, nous présentons un modèle formel de graphe spatio-temporel issu d'un modèle de la littérature, que nous adaptons afin de prendre en compte les spécificités des données géohistoriques. Ce *graphe géohistorique* nous permet de représenter les transformations existantes entre les observations d'une base de données spatiale et temporelle. Pour identifier ces transformations, nous nous intéressons ensuite à la notion d'identité dans le contexte des données spatio-temporelles, et nous proposons un cadre pour définir de façon souple cette identité afin de spécialiser un graphe géohistorique à un cas d'application précis. Nous proposons ensuite un processus de construction d'un tel graphe à partir d'un ensemble d'observations géohistoriques structurées selon notre modèle de données spatial et temporel. Nous introduisons pour cela une nouvelle structure, appelée *hypergraphe de processus*, modélisant les processus de transformation de l'espace et servant de base à notre méthode de construction.

Enfin, nous proposons dans le **chapitre 6** une méthode de construction semi-automatique d'un graphe géohistorique à partir d'une définition de l'identité des observations. Pour cela, nous mettons en place une procédure en deux temps. D'abord, nous construisons par optimisation stochastique l'hypergraphe de processus d'un ensemble d'observations géohistoriques. En identifiant les observations dont les identités sont liées¹, nous découvrons les relations de filiation qui lient les observations entre elles. Ensuite, nous présentons une méthode fondée sur la théorie des fonctions de croyance permettant, à partir d'une certaine définition de l'identité, d'identifier la nature exacte des relations de filiation préalablement construites. Ainsi, nous créons un graphe géohistorique complet à partir d'observations éparpillées stockées au sein de *snapshots* dans la base de données spatiale et temporelle.

1. Parce qu'elles représentent une même entité du monde réel, ou deux entités dont l'une est le résultat de la transformation de l'autre.

Première partie

Positionnement et sources
cartographiques sur Paris

Chapitre 1

Positionnement du sujet

Objectifs :

- Explorer l'adéquation entre l'utilisation de l'espace par les historiens et les fonctionnalités offertes par les SIG historiques existants.
- Identifier les complexités des données géohistoriques.
- Formaliser la problématique de la thèse et se positionner par rapport à elle.

Sommaire

1.1	Analyser l'espace pour comprendre l'évolution d'une société urbaine	14
1.1.1	Le choix de l'espace	14
1.1.2	Quels usages de l'espace urbain en histoire?	16
1.1.3	Conclusion : besoins généraux	22
1.2	Représenter un espace urbain en transformation	23
1.2.1	Information géohistorique	25
1.2.2	Typologie des sources d'information historiques sur l'espace urbain	27
1.2.3	Le plan de ville au cœur de la représentation visuelle d'un espace urbain ancien	32
1.3	L'apport des SIG pour étudier l'histoire : un choix méthodologique	38
1.3.1	Introduction : définitions préalables	38
1.3.2	Représentation de l'information historique	41
1.4	Les difficultés dues aux imperfections de l'information géohistorique	44
1.4.1	Le processus d'intégration d'une information géohistorique	46
1.4.2	Types d'imperfections de l'information géohistorique	48
1.5	État de l'art des SIG utilisés en histoire	50
1.6	Positionnement	57
1.7	Conclusion	61

Le passé ne se conserve pas, mais fait l'objet d'une reconstruction toujours recommencée.

Carnets de croquis. Sur la connaissance historique , Bernard Lepetit, 1999.

Ce chapitre propose d'explorer plus avant les raisons de l'inadéquation entre des outils informatiques, pourtant destinés à gérer des données spatiales, et la pratique historique d'un espace urbain. Après avoir formalisé et défini les particularités de l'information géographique utilisée en histoire -désignée alors par le terme *géohistorique*- et exposé les différentes sources d'information géographiques utilisées dans le cadre de recherches historiques sur l'espace urbain, nous nous attacherons à décrire le processus d'intégration d'une information au sein d'un SIG historique afin d'identifier les différentes sources d'imperfections dans les sources et l'information utilisée. Enfin, nous présenterons différents travaux existants visant à manipuler une information géohistorique imparfaite dans le cadre des systèmes d'information géographiques.

D'abord, nous regarderons du côté des utilisations de l'espace urbain en histoire afin d'identifier les besoins et attentes spécifiques des historiens. Pour ce faire, nous commencerons dans la section 1.1 à mettre en lumière les pratiques des historiens et leurs problématiques grâce à l'analyse d'un espace urbain. Dans la section 1.2.2, nous présenterons une typologie des différentes ressources permettant de reconstruire une représentation d'un espace urbain ancien. Il s'agira non seulement d'explorer les différentes spécificités de ces ressources, mais également d'en identifier les complexités. Ainsi, nous pourrions expliciter clairement les objectifs des historiens quant aux usages de l'espace urbain au sein de leurs études, et par là identifier les demandes adressées par l'étude historique à des outils permettant de manipuler et représenter des données géographiques.

Dans un second temps, nous nous intéresserons plus précisément aux systèmes d'informations géographiques et à leurs usages pour des études historiques. Nous aborderons dans la section 1.3 le regard des historiens sur les SIG historiques à partir de la littérature de façon à classifier les différents objectifs et problèmes posés à l'encontre des SIG par les chercheurs en histoire. Cette section permettra également de terminer la description du processus de reconstruction de l'espace et d'établir, dans la section 1.4, une typologie des différentes imperfections de la représentation obtenue liées aux sources, aux données et au traitement d'une information historique et spatiale - géohistorique- au sein d'un SIG.

Finalement, nous croiserons ces deux regards pour analyser les systèmes d'informations géographiques destinés à l'histoire existants dans la littérature. Ainsi, dans la section 1.5, les propositions existantes seront confrontées aux attentes et aux complexités identifiées dans les sections précédentes. De cette façon, nous montrerons les points d'adéquation, mais également les manques des propositions existantes. En particulier, nous nous positionnerons (section 1.6) par rapport aux problématiques

d'intégration de l'information historique issue des représentations anciennes de l'espace urbain, et de modélisation des dynamiques des structures composant cet espace.

1.1 Analyser l'espace pour comprendre l'évolution d'une société urbaine

1.1.1 Le choix de l'espace

Éléments de définition

Dans le travail présenté ici, l'espace occupe une place prépondérante. Celui de la ville, dans son organisation et ses transformations au cours du temps constitue le vecteur d'intégration de l'information historique. Étant donné ce statut central, il nous paraît essentiel de fixer les frontières de ce terme dans le cadre d'un travail se situant quelque part entre géomatique et histoire. En effet, l'espace englobe des réalités variées, liées aux sensibilités des disciplines qui manipulent ce concept, allant de l'espace physique, isotrope et homogène à un espace totalement lié aux hommes, produit des sociétés qui le construisent (Lefebvre, 2000). L'espace est en premier lieu une notion géographique que Pumain et Saint-Julien (Pumain, 1993; Pumain et Saint-Julien, 1997) déclinent sur deux axes :

- Soit il s'agit de l'espace physique, *ensemble de lieux (ou objets localisés) muni d'une distance entre ces lieux*. Il est le *simple contenant des objets* qui permet d'analyser leurs relations.
- Soit il s'agit d'un *ensemble de lieux et de relations entre eux, définies par les interactions entre des acteurs sociaux localisés*. C'est le *produit de l'organisation et de la nature, agent du maintien et du développement des sociétés sur leur territoire*.

On parle alors généralement d'espace **support** et d'espace **produit**, le premier étant vu comme le **socle immuable sur lequel se situent les faits sociaux**, le second étant pleinement **modelé par les activités des sociétés qui l'occupent** (Méo, 1991). Des nombreux travaux, tels ceux, fondateurs, de Roncayolo et Chesneau (2011) et Lepetit (1980), abordent les relations entre histoire et géographie et dressent, pour les sciences sociales, le portrait d'un espace intrinsèquement lié aux sociétés, produit d'une construction constante et renouvelée¹. Partant de cet angle de vue, notre espace est donc bien physique, topographique, mais impliqué dans une relation avec les phénomènes sociaux qui l'inscrivent fortement dans la dimension temporelle. Dans tous le reste du document, le mot "espace" fera référence à cet espace produit. Quant à l'espace urbain, nous choisissons de le délimiter par la frontière administrative de la ville, élargie éventuellement à ses abords immédiats. Dans le Paris ancien, il s'agit de l'étendue à l'intérieur de l'enceinte des Fermiers Généraux, à laquelle s'ajoute les communes annexées en 1860². Cette articulation profonde entre espace géographique et temporalité se trouve notamment à la racine

1. Le mouvement au sein des sciences humaines qui donne à l'espace une rôle majeur en l'utilisant comme outil de compréhension des phénomènes sociaux est désigné par l'expression *spatial turn*. Les spécificités de cette tendance ont été abordée notamment par Cosgrove (1984), ou plus récemment Torre *et al.* (2008).

2. On dévie ici légèrement de la définition habituelle où l'espace urbain repose sur la continuité du bâti (Moriconi-Ebrard, 1993). Dans le Paris ancien, cette continuité n'existe que dans un espace central restreint, où elle est rapidement brisée par les espaces vides ou cultivés à l'intérieur même des enceintes.

des travaux de Braudel (Bataillon, 1950; Braudel, 1951), qui plaide pour une étude des phénomènes sociaux inscrite dans un espace géographique en transformation (ce qu'il nommera la *géohistoire*). Il considère trois couches imbriquées de temporalités qui constituent l'histoire : la première est celle des individus, rapide et nerveuse, c'est *une agitation de surface*. La seconde temporalité est plus lentement rythmée, c'est celle des groupements, des civilisations, de l'économie et des sociétés. La dernière, pratiquement figée, correspond aux rapports des Hommes avec le milieu qui les entourent, avec l'espace support de leur établissement. L'espace de l'histoire, produit des sociétés, est donc inscrit dans des temporalités aux rythmes multiples. S'il nous faut préciser les formes de l'espace urbain qui nous occupent ici, il nous faut dans le même temps définir les temporalités particulières que nous considérons.

Intégrer le temps pour étudier la ville

Pour Lepetit, reprenant Braudel (Lepetit, 1986), l'espace géographique est une dimension supplémentaire pour l'histoire, constitué de *valeurs dormantes* (par exemple les axes de circulations marchands) s'offrant à l'analyse historique. La ville, au travers de ses structures, de son organisation, est alors un *conservatoire temporel* (Lepetit, 1986) offrant à l'historien des fragments de ses configurations passées. Cette idée d'une ville constituée d'une accumulation de temporalités passées, dans la lignée de Roger Dion (Courville, 1995)³ sera de nouveau exploitée par Lepetit (1996), pour qui l'histoire intègre par la dimension temporelle les considérations sociales, politiques, économiques sur la ville. L'étude de la ville en histoire -l'histoire urbaine- est alors un moyen de compréhension de l'évolution d'une société au travers de son étude sur le temps long⁴. Perrot (1975) invoque également cette inscription de la ville dans le temps -pour l'Europe tout du moins-, arguant que la compréhension de l'urbanisation et de ses formes repose sur des processus de transformations urbaines dans la longue durée. L'histoire urbaine inscrit donc l'espace de la ville dans le temps pour en étudier les transformations et les mettre en rapport avec les phénomènes sociaux qui s'y sont déroulés.

Pour Roncayolo (d'après Walter (2005)), *les pierres et les formes urbaines sont du temps et des pratiques consolidées*. Attaché à faire apparaître les divisions sociales au travers de ses *lectures de villes*, il y décrypte les relations entretenues par les portions d'espace et les formations sociales. Ces relations se lisent alors dans la construction de l'espace urbain, c'est-à-dire son inscription dans différentes *couches temporelles* aux rythmes variés -le temps devient alors le *tempo-*, aboutissant à des décalages, des ruptures qui se lisent dans cette construction urbaine. Son étude des limites de Marseille et Paris aux XIX et XXe siècles, d'abord matérialisées par des enceintes puis persistant au travers de limites symboliques, s'appuie à la fois sur l'organisation et les transformations des structures urbaines mais aussi sur des

3. *Tout paysage humanisé est le reflet d'une histoire, [...] un amalgame d'apports inégaux en âge*. Roger Dion, d'après (Courville, 1995, p. 20)

4. L'archéologie fait un large usage de ce terme pour désigner des périodes de plusieurs siècles ou millénaires (Redman et Kinzig, 2003). En histoire, les périodes d'étude sont souvent plus restreintes. Nous désignons par le terme "temps long" toute durée suffisamment grande pour que des changements sensibles dans l'organisation de l'espace se produisent. Dans l'espace parisien du XIX^e siècle, quelques dizaines d'années suffisent pour modifier en profondeur les réseaux de rues et l'organisation du bâti

considérations sociales (contraintes militaires, opérations immobilières et financières, etc.) et symboliques (représentation mentales de la ville et de ses limites) (Walter, 2005).

D'après ces historiens, il s'agit donc d'étudier les structures de la ville, son organisation, des réseaux de voies de communication aux transformations des structures bâties qui la composent. Pourtant, comme le fait remarquer Lemas (2009), le champ de l'histoire urbaine n'est pas homogène et l'importance prise par l'espace urbain dans l'étude varie largement. Cette hétérogénéité se traduit notamment par des degrés d'importance attribués aux dynamiques de l'espace dans l'étude historique, conditionnant alors la place prise par la dimension spatiale dans l'étude. Nous faisons ici l'hypothèse que ces différents usages vont conditionner les attentes des historiens, et donc les outils et les méthodes appropriés. Il est alors nécessaire d'identifier les différents usages de l'espace de la ville en histoire de façon à nous positionner efficacement par rapport aux différents besoins des historiens, et à pouvoir analyser l'adéquation des outils existants avec ces besoins.

1.1.2 Quels usages de l'espace urbain en histoire ?

Dans le paragraphe précédent, nous avons expliqué que l'espace géographique en histoire était toujours traité comme un produit résultant des activités humaines. Pourtant, nous allons le voir, la place occupée par l'espace de la ville est très variable au sein de l'historiographie. Nous proposons alors d'explorer cette variété au travers de trois modalités fondées sur l'importance de la relation entre l'espace étudié et le temps. Nous choisissons volontairement de ne pas considérer les phénomènes sociaux étudiés comme variable de contrôle car leur variété et leur hétérogénéité empêcheraient tout regroupement. Les dynamiques spatiales se mêlent d'autant plus aux dynamiques sociales étudiées que la liaison entre espace et temps augmente. Ce couplage a cependant un coût : il implique un faible degré d'abstraction de l'espace et une finesse temporelle importante. Il conditionne donc les besoins des historiens en termes de modèles et de représentations de l'espace. Dans les paragraphes suivants, nous présentons ces différents niveaux d'intégration du temps et identifions les besoins qui leur sont spécifiques à partir de différents travaux en histoire. Bien sûr, les exemples abordés ne constituent qu'une infime partie de la production en histoire urbaine, cependant nous considérons qu'ils illustrent correctement le propos.

L'espace urbain : cadre de l'étude

Mettre en relation un phénomène social avec un espace géographique constitue le premier niveau de prise en compte de l'espace en histoire. Cette mise en relation implique de situer dans l'espace les différents phénomènes et faits étudiés, construisant ainsi le contexte géographique de l'étude historique. À la manière d'un fond cartographique servant de base à la représentation graphique d'une information géographique, l'espace sous-tend et illustre le propos historique, mais son organisation n'en constitue un élément d'explication qu'à un niveau d'abstraction et une échelle élevés.

Dans son travail sur le cosmopolitisme dans les villes méditerranéennes et plus particulièrement celles de l'empire Ottoman, Lafi (2011) explore les différentes formes de ce phénomène, les rapports de pouvoirs entre communautés au travers du prisme des volontés politiques des différents régimes (empire Ottoman et royaumes occidentaux notamment), des valeurs portées par les religions et des traits culturels locaux. Ainsi, si la ville se trouve au cœur de cette étude, son espace n'est que très peu considéré. En effet, l'évocation des villes par leurs noms fournit surtout une toile de fond sur laquelle est dessinée l'étude, servant à différencier les formes de cosmopolitisme par la ville et sa position dans de grandes zones géographiques (par exemple le nord de la Méditerranée). Si le phénomène s'inscrit dans le temps tout au long de l'étude, l'espace quant à lui est figé. Il reste très abstrait, évoqué uniquement par des noms de lieux, d'aires culturelles et constitue, à proprement parler, le support de l'étude. L'organisation interne des villes n'est alors pas objet d'étude, bien qu'elle se lise par exemple en filigrane dans l'évocation de l'identité communautaire des quartiers (par exemple à Tripoli). Cet espace n'est pourtant pas totalement neutre. En effet, la situation portuaire de certaines villes est développée pour expliquer certains profils de cosmopolitisme.

Dans son travail sur la notion *d'habiter* dans les villes italiennes, Barbot (2013) s'inscrit plus directement dans l'espace urbain, étudiant pour plusieurs villes d'Italie les pratiques sociales liées au fait d'habiter la ville et les rapports de forces entre les couches de la société. Plus particulièrement, l'auteur montre parfaitement la conséquence de ces pratiques sur l'espace, les différentes couches sociales présentes dans la ville étant visibles au travers des types d'habitats correspondant à leur degré d'intégration dans la ville. Cet espace constitue un cadre dans lequel viennent se placer les différentes formes de l'appartenance à la ville, mais son organisation interne ou sa situation géographique ne sont pas des facteurs de compréhension de ces appartenances. La ville reste un objet abstrait, spectateur d'une étude dont le sujet porte essentiellement sur les rapports sociaux.

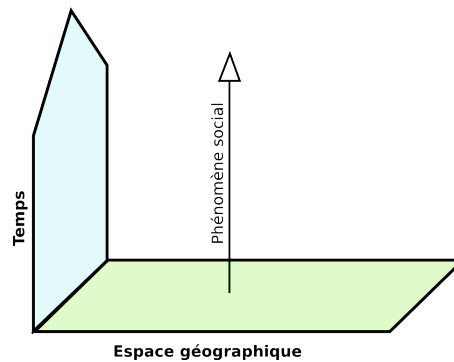


FIGURE 1.1 – Premier niveau d'intégration de l'espace urbain : c'est un cadre servant de contexte, implicitement situé dans le temps par le thème de l'étude. La dimension temporelle est dominante. **T**

Au travers de ces deux travaux, nous pouvons voir apparaître un premier niveau où l'intrication du temps et de l'espace urbain est très peu présente. Si le phénomène social s'inscrit toujours dans le temps, l'espace urbain est quant à lui figé dans une temporalité longue. Il constitue le contexte du propos plutôt qu'un objet d'étude à part entière (voir figure 1.1). L'étude n'est pas spatiale, mais fait référence à l'espace, elle reste purement temporelle. En marge de l'étude, il peut être

abordé de façon très abstraite (Arnaud, 2008), ses formes apparaissant en filigrane par l'évocation des quartiers, des séparations entre groupes sociaux ou des positions relatives des objets abordés. Dans les deux cas, l'échelle est relativement grande, permettant d'englober tout l'espace concerné par le phénomène étudié.

L'espace urbain outil d'étude

D'autres auteurs trouvent dans l'organisation de l'espace un outil de compréhension des phénomènes qu'ils étudient. Il s'agit alors la plupart du temps d'aborder l'espace comme un support sur lequel l'on vient localiser des données sociales dans le but de faire apparaître des configurations particulières et des relations jusqu'ici invisibles. Il s'agit finalement de considérer que *la distribution des phénomènes dans l'espace n'est pas neutre dans la manière dont ils évoluent et se transforment* (Arnaud, 2008). Il ne s'agit donc pas encore d'étudier l'organisation et les formes de l'espace en elles-mêmes, mais de lire au travers d'elles des propriétés, des relations entre les phénomènes étudiés que seule leur situation dans l'espace de la ville permet de faire apparaître efficacement (citons par exemple les travaux de Marraud (2010) sur les corporations parisiennes aux XVII^e et XVIII^e siècles, appuyant dans l'espace de Paris son analyse des résistances et des transformations de la corporation des boulangers au gré des rapports de forces qu'elle entretient avec le pouvoir royal et de l'ordre public).

L'espace géographique est aussi un outil de compréhension des phénomènes sociaux. Inscrits dans celui-ci, leur répartition et organisation offre de nouvelles pistes de réflexion et d'analyse. Il s'agit en quelque sorte de mettre les phénomènes à l'épreuve de l'espace. On peut en particulier citer les travaux de Gauthiez et Zeller (2010) qui restituent à partir de textes d'archives les trajets des milices bourgeoises au XVII^e siècle dans un parcellaire de la ville de Lyon, permettant alors de mettre en évidence différents profils de parcours et faisant naître de nouvelles questions liées aux formes de ces parcours et à leurs transformations. Weiss (2009) dans ses travaux sur les censives parisiennes inscrit directement les rapports sociaux entre propriétaires parisiens. Les divers conflits juridiques se transcrivent dans l'espace de la ville par des zones où l'appartenance à une censive⁵ est mal définie.

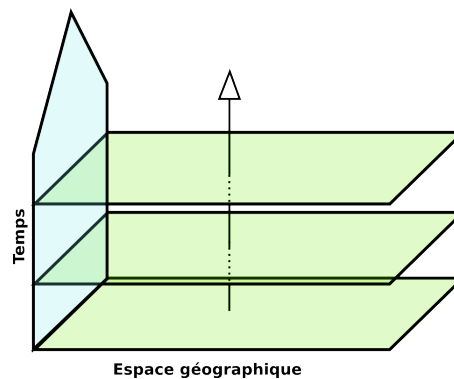


FIGURE 1.2 – Second niveau d'intégration de l'espace urbain : l'espace devient un outil d'analyse des données sociales permettant de les répartir dans la ville. Les dimensions spatiales et temporelles sont considérées conjointement (S+T)

5. La censive est une terre appartenant à un seigneur, dont il cède certains droits en échange d'une redevance : le cens.

Ces études⁶ s'appuient généralement sur des méthodes d'analyse spatiale, en particulier sur des représentations cartographiques des phénomènes dans lesquelles des données qualitatives (par exemples les surfaces des censives) ou quantitatives discrétisées (statistiques notamment) sont rendues visibles. La carte est un puissant outil d'interrogation d'un espace passé, fournissant une vue globale de l'organisation spatiale des données, mais également un outil permettant de remettre en question les faits historiques analysés (Arnaud, 2008).

Cette fois l'espace urbain est considéré, à la manière de Lepetit (1996), comme un témoin des activités humaines, une construction qui se donne à lire et permet d'extraire une connaissance sur un fait historique. L'espace n'est plus abstrait, il est reconstruit à partir des traces qui en subsistent - plans, textes, etc.-, l'étude étant tributaire des sources de descriptions existantes. Cet espace reconstruit est alors situé dans le temps, mais il s'agit encore d'un socle d'étude, ses transformations ne font pas l'objet d'une attention particulière. Lorsque le phénomène social évolue dans le temps, l'espace doit être reconstruit de nouveau à partir d'autres sources pour que sa forme et son contenu soient contemporains du phénomène. La ville est alors traitée comme une accumulation de tranches temporelles, comme illustré dans la figure 1.2. L'étude se place cette fois dans un cadre à la fois spatial et temporel.

L'espace urbain : objet d'étude

Si l'on considère que l'espace urbain est un produit des activités humaines, son organisation et ses structures gardent la mémoire des activités humaines passées (C'est ce qu'Arnaud (Arnaud, 2008, p. 141) nomme *hypothèse de permanence*). Il devient donc un outil de lecture au travers duquel il est possible de recomposer une histoire urbaine. Ces traces se trouvent accumulées dans l'espace, transcrites notamment par les brisures ou au contraire les persistances des structures urbaines (Ducom, 2009). L'enchevêtrement qui en résulte doit être décrypté, en *déroulant* l'histoire de la ville, à partir des structures persistantes, des connaissances des structures passées et des événements ayant affecté telle ou telle structure. On peut citer en exemple les travaux de Boudon et Blécon (1977), analysant dans le détail les transformations du quartier des Halles à Paris dans le temps long.

L'objectif n'est pas seulement de retracer les transformations de la ville, mais de considérer qu'il existe une boucle de rétroaction entre l'espace géographique et les phénomènes sociaux, lisible uniquement dans le temps. Ainsi, lorsqu'il étudie les évolutions de l'espace social dans un îlot parisien aux XVIII^e et XIX^e siècles (voir figure 1.3), (Gribaudo, 2009) place l'organisation spatiale de l'îlot au cœur de l'étude. Ancien hôpital nationalisé sous la Révolution, l'îlot se transforme petit à petit en un espace dense et vivant composé d'ateliers d'artisans et de commerçants. Ces transformations, loin d'être homogènes dans leurs localisations et leurs rythmes sont abordées comme le produit de pratiques locales (la spéculation immobilière) et globales (la vente des biens nationaux⁷). Inversement, les pratiques sociales sont

6. On peut encore citer Gonzalez-Quijano (2012) ou bien Montel (2013)

7. Les biens nationaux correspondent aux biens de l'Église et les domaines seigneuriaux nationalisés à la suite du décret du 2 novembre 1789.

mises en rapport avec la perception de l'espace îlot au travers des perceptions des habitants, mais aussi des administrateurs et observateurs extérieurs. D'autres auteurs effectuent le même travail de mise en relation, par exemple pour étudier les évolutions du prix du foncier selon les restructurations du parcellaire parisien (Bove, 2012), ou bien encore pour retracer l'évolution des structures urbaines (parcellaire et voirie) d'un espace local à Tours (Lefebvre, 2009).

Cette fois l'organisation de l'espace urbain lui même est analysée, soit indépendamment de données sociales particulières (l'espace est alors considéré comme gardant la mémoire des activités qui l'ont modelé (Roncayolo et Chesneau, 2011)), soit en relation avec un ou plusieurs phénomènes sociaux. Lorsqu'il s'agit de l'organisation de l'espace seul, l'analyse morphologique via la représentation cartographique des espaces anciens en constitue le principal outil d'exploration (Arnaud, 2008). Insérée dans le temps, l'analyse morphologique doit également s'intéresser aux changements qui affectent les formes urbaines. Elle peut être visuelle, par la mise en correspondance de cartes du même espace à des instants différents (c'est le cas par exemple de l'espace étudié par Lefebvre (2009)), ou bien s'appuyer sur des indicateurs morphologiques (Strano *et al.* (2012) ou Barthelemy *et al.* (2013) par exemple) qui étudient les transformations des réseaux de rues dans la région de Milan, puis à Paris via les d'indicateurs structurels sur ces réseaux). Finalement, lorsque l'étude porte sur les relations entre le déroulement des phénomènes sociaux et les transformations de l'espace urbain, l'analyse morphologique vient compléter l'analyse spatiale.



FIGURE 1.3 – Transformations des structures morphologiques de l'îlot de la Trinité à Paris sur trois périodes : au milieu du XVIII^e, en 1845 et après les percées haussmanniennes. Illustration extraite de Gribaudo (2009).

Recomposer un espace dans toute la complexité de ses transformations implique un important volume d'informations décrivant les structures urbaines, leur organisation et les événements qui les affectent. Par conséquent, les espaces traités sont souvent restreints mais font l'objet d'une description extrêmement précise. L'espace est alors intimement lié au temps, en interaction continue avec les phénomènes sociaux qui y prennent place. L'étude n'est plus seulement spatiale et/ou temporelle car les deux dimensions sont indissociables : elle devient **spatio-temporelle**.

L'espace urbain : un système (complexe)

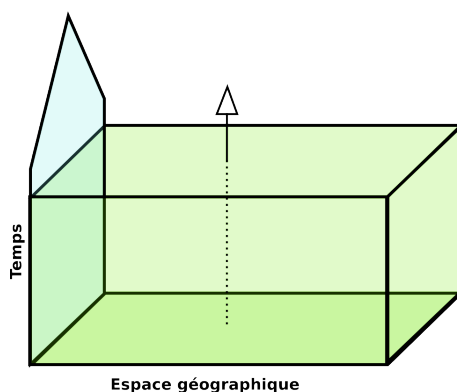
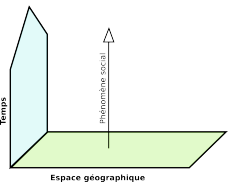
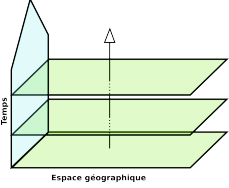
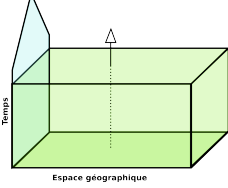


FIGURE 1.4 – Troisième niveau d'intégration de l'espace urbain : l'espace en transformation constante est mis en rapport avec les phénomènes sociaux étudiés, chacun améliorant la compréhension de l'autre. L'espace devient dynamique, l'analyse spatio-temporelle (**ST**)

Le tableau 1.1 résume les trois niveaux de prise en compte de l'espace et de son imbrication avec le temps en histoire. On constate au travers des exemples présentés précédemment que l'échelle et le niveau de d'abstraction spatiale varient. Ceci s'explique notamment par le fait que lorsque l'espace n'occupe qu'un statut de cadre, son détail n'est pas nécessaire à l'analyse mais des études sur de larges étendues peuvent être menées. Lorsque les structures mêmes de l'espace interviennent dans l'étude, celle-ci se trouve contrainte par la complexité liée à la reconstruction d'un espace ancien : quantité de sources d'archives nécessaires, difficultés de traitement de ces sources, etc. Le volume de données nécessaires sur l'espace augmente encore lorsque l'on s'intéresse aux transformations de l'espace. L'étude nécessite non seulement une connaissance sur les formes de la ville mais aussi sur les événements qui les affectent, réduisant encore son échelle spatiale considérée.

Le parcours des différentes considérations spatiales dans les travaux historiques ne fait finalement pas ressortir des catégories séparées, mais plutôt un échelonnement dans la part attribuée non pas seulement à l'espace de la ville, mais surtout à son insertion dans le temps et aux rapports qu'il entretient avec les phénomènes sociaux qui s'y déroulent. Finalement, l'espace géographique totalement inclus dans le temps porte à la fois la marque des actions humaines, mais son impact sur les pratiques et la perception par les individus fait également l'objet d'une attention particulière : il s'agit de considérer l'espace dans un système *au sein duquel il fonctionne selon une boucle de rétroaction avec la société, qui aménage, gère et organise le territoire, tandis que le territoire rétro-agit sur la société* (Raffestin, 1986). Dans cette relation systémique, les objets géographiques qui composent l'espace urbain (bâtiments, rues, parcelles, etc.) doivent être considérés comme des objets dynamiques, en transformation dans le temps. Ils ne sont plus seulement spatiaux et insérés dans le temps à une date ou une période donnée mais ils intègrent cette dimension pour devenir pleinement **spatio-temporels**.

TABLE 1.1 – Résumé des trois niveaux de relations entre espace urbain et temps et des propriétés particulières de chaque niveau.

	Cadre	Outil	Objet
			
Abstraction spatiale	+++	+	+
Échelle⁸	macro (ensemble de villes, pays, etc.)	méso (ville)	méso ou micro (quartier, bâtiment, etc.)
Dimensions	T	S+T	ST

1.1.3 Conclusion : besoins généraux

Les 3 catégories d'intégration du temps dans l'étude de l'espace urbain ne sont finalement pas isolées les unes des autres, mais résultent d'une gradation de l'insertion de l'espace urbain dans le temps. De cette imbrication naît la relation entre l'espace et des phénomènes sociaux particuliers ; c'est considérer que l'espace de la ville permet de mieux comprendre les phénomènes sociaux qui s'y sont déroulés parce qu'il influe sur ces activités et qu'il en porte les traces résiduelles. Un tel gradient fait écho à la volonté de Roncayolo, illustrée par Walter (2005), d'étudier la ville en effectuant une *réduction au temps* de la même manière que Braudel (1951) plaidait pour une *réduction à l'espace*.

Spatialiser des informations sociales se concrétise la plupart du temps au travers de la production de cartes (Arnaud, 2008, p. 102-103). Manipuler l'espace ancien signifie donc le représenter : mentalement d'abord, mais aussi visuellement par la construction de cartes, la mise en commun et le croisement de sources historiques. De cet intérêt pour la construction de représentations du passé, on peut dégager trois besoins particuliers :

1. **Localiser des données sociales** décrivant un des phénomènes ou des faits sociaux, dans l'espace urbain **contemporain** de leur époque. Abstrait, l'espace reste le contexte de l'étude mais il permet d'expliquer certaines spécificités d'un phénomène par sa localisation générale (par exemple les particularités du cosmopolitisme dans les villes côtières chez Lafi (2011)). Lorsqu'il est plus concret, l'espace permet de situer les données sociales et de faire apparaître des motifs d'organisations des individus, des jeux de pouvoirs qui se traduisent dans la répartition des individus, des événements, etc.
2. **Représenter l'organisation des structures de l'espace et ses transformations au cours du temps**, dans le but de faire apparaître des motifs,

des brisures ou à l'inverse des zones de résiliences de l'espace.

- 3. Représenter l'évolution conjointe de l'espace et d'un phénomène social qui s'y déroule.** Il s'agit d'un besoin lié au niveau de relation espace/temps le plus fort, dans lequel les relations de rétroaction entre l'espace urbain et les activités sociales sont directement considérées au sein de l'étude.

Plus dynamiques spatiales et sociales sont mêlées, plus il devient important de reconstituer les états et les transformations de l'espace afin d'en produire une représentation visuelle. La complexité de ces transformations et le niveau de détail demandé par de tels travaux impliquent une connaissance fine de l'espace étudié ne pouvant être acquise qu'au travers des traces qu'il faut rassembler et lier au sein d'une représentation visuelle cohérente.

1.2 Représenter un espace urbain en transformation

Représentation cartographique de l'espace

Nous avons présenté la demande d'une représentation de l'espace passé pour l'histoire urbaine. Il peut en effet sembler naturel qu'un travail sur l'espace implique la production de représentations visuelles de celui-ci, notamment cartographiques, étant donné la force du lien ainsi créé avec la géographie. La représentation de l'espace, comme le fait remarquer Palsky (2004), est une notion relativement complexe et surtout doublement connotée : c'est à la fois le processus de création d'une image mentale mais aussi l'image elle-même. La carte est une concrétisation de cette représentation sous forme d'image stabilisée, c'est *une abstraction de la réalité spatiale [...] modélisée et codée afin d'être appréhendée par le regard*.

Lorsqu'il s'agit de représenter un espace passé, le processus de représentation cartographique peut intervenir à deux moments distincts. En effet, l'espace dont on souhaite produire une représentation n'existe plus et n'est donc plus accessible à l'observation directe. En conséquence, il doit d'abord faire l'objet d'une reconstruction à partir des traces qui se sont décantées au cours du temps⁹. C'est ce point précisément que nous allons développer par la suite dans la section 1.2.2. Une fois cette représentation construite, elle peut servir de base à l'analyse ou d'outil permettant de localiser des informations diverses. À partir du moment où l'on dispose d'une représentation visuelle d'un espace ancien, elle peut servir à la communication des résultats du travail ou bien de contexte visuel pour la représentation et l'interprétation de données sociales. C'est donc un processus en deux temps qui s'impose lorsqu'il s'agit de travailler sur une configuration spatiale passée. Nous proposons dès lors d'analyser plus avant ce processus pour en identifier les complexités, en présentant d'abord les différents types de traces considérées sur l'espace. Nous nous concentrerons dans la section 1.2.3 sur les sources cartographiques, qui par leur na-

9. Une trace est la marque d'un objet ou d'un événement dont la particularité est *d'avoir été*. Une trace crée donc systématiquement un décalage entre le moment de sa création et sa lecture, décalage pendant lequel elle peut se décomposer ou se désintégrer (Krämer (2010)). Les restants d'un ancien mur, un récit de voyage ou un portulan sont autant d'exemples de traces d'un espace passé.

ture se prêtent le plus naturellement à la construction de représentations visuelles de l'espace.

La recherche de traces

Dans son livre *History and GIS*, von Lünen (2013) fait le parallèle, à titre illustratif, entre le travail d'un historien et ceux d'un enquêteur ou d'un chasseur. Là où l'enquêteur cherche des pièces à conviction lui permettant d'établir la culpabilité ou l'innocence d'un accusé, ou encore de chercher le coupable d'un crime, l'historien cherche au sein d'archives les éléments lui permettant de répondre à une question historique, de retracer les faits, les événements et leur contexte historique. Lünen note cependant que les méthodes sont radicalement différentes. À la différence de l'enquêteur qui se base sur des procédés de déductions lui permettant de recomposer une scène de crime, l'historien procède par abduction, c'est à dire par la formulation d'une hypothèse historique qu'il va tenter de faire apparaître au travers de l'agencement des traces historiques qu'il a à sa disposition.

Cette analogie peut sembler grossière, mais elle met assez bien en évidence la façon de procéder d'un historien : à partir d'une question historique précise, il va chercher des faits tangibles lui permettant d'y répondre. Il trouve dans le particulier des éléments lui permettant, par généralisation, de répondre à une question plus large. Lorsque cette question est profondément spatiale, ou bien que l'historien fait l'hypothèse que l'espace y joue un rôle, le travail historique passe par la mise en relation de traces spatiales ou spatialisables issues de sources historiques diverses.

Au delà même des traces directement lisibles au sein de sources historiques, leur croisement permet de faire apparaître de nouvelles informations. Par exemple, la comparaison de deux cartes d'un quartier à des époques différentes peut permettre d'acquérir des connaissances sur la façon dont le quartier s'est transformé. Du fait de l'importance de ce travail de recherche de traces, de croisement et d'extraction de connaissances nouvelles, il est crucial de garder en tête ses particularités lors de la création d'outils et de modèles SIG adaptés au travail historique.

Recomposer n'est pas répliquer

Lorsque l'on crée une représentation d'un espace ancien, il ne n'agit pas d'essayer de reproduire la réalité à un temps donné, d'autant plus qu'il est impossible d'appréhender une réalité complexe. Cet état de fait est largement admis dans les sciences sociales, mais également parmi les disciplines traitant de l'information (géographique notamment) (Bodenhamer *et al.*, 2010). Pour la géomatique en particulier, cette question se trouve à la source des modèles de qualité des données et d'imperfections de l'information spatiale (Leyk *et al.*, 2005). L'historien, lui, manipule des sources -des traces- dont il connaît les défauts et dont il questionne systématiquement la fiabilité. Il faut s'interroger également sur la capacité des SIG à gérer de telles sources et à manipuler des informations dont la fiabilité est très variable.

1.2.1 Information géohistorique

Dans le domaine des systèmes d'informations géographiques, on désigne par "information géographique" le couple formé d'une information portant sur un objet ou un phénomène du monde terrestre et d'une localisation à la surface de la Terre dans un système de référence explicite (Denègre et Salgé, 1996). La perception de cette information géographique est classiquement considérée comme le fait de structurer l'espace réel sous la forme d'entités, correspondant à un phénomène quelconque du monde réel (une personne, un bâtiment, l'Europe, un champs de bataille, etc.) (Plewe, 2002). Le champ de bataille et la personne sont pourtant des entités très différentes. Leurs limites en particulier n'ont rien de comparable. Dans le cas d'une personne, à l'échelle macroscopique du moins, sa limite est établie *de facto*. Il en va d'une manière tout à fait différente pour le champs de bataille : ses limites résultent de processus cognitifs plus complexes.

Entités

Cette différence entre entités géographiques a été formalisée par Smith (1995). Il classifie ainsi ces différentes entités en deux groupes de natures différentes :

- Les entités *bona fide* présentent des limites directement perceptibles. À une échelle macroscopique tout du moins, une personne ou une île sont des entités *bona fide*. Leurs limites sont clairement visibles et la correspondance entre l'objet du monde réel "personne" ou "île" et sa représentation cognitive est directe.
- Les entités *fiat* sont au contraire des entités dont les limites sont créées de toutes pièces par l'Homme. Un pays, une paroisse en sont des exemples : leurs limites peuvent localement s'appuyer sur des éléments naturels, mais elles restent des constructions humaines.

Plewe (1997) ajoute un troisième groupe qu'il nomme *motivated entities*. Comme les entités *fiat*, leurs limites n'existent pas de façon concrète dans la réalité, mais cette fois elles résultent de processus cognitifs plus ou moins complexes tels l'agrégation, la catégorisation ou la simplification. La ville est clairement une entité motivée. Les délimitations d'une ville procèdent de regroupements de structures humaines dont la densité ou l'organisation définissent une certaine notion d'urbanité.

Smith et Varzi (2010) font cependant remarquer que la séparation entre entités *fiat* et *bona fide* n'est pas si évidente : les limites d'une entité peuvent être mélangées. Ainsi, les quartiers parisiens actuels sont des entités *bona fide* dont les contours sont définis par des rues ou la Seine, mais qui sont également des constructions humaines sans existence *de facto*. Les entités sont alors moins définies par leurs frontières seules que par opposition à d'autres entités. Nous pensons que la même critique peut être faite pour les entités motivées. Les faubourgs anciens de Paris se définissent dans leur rapport d'opposition avec la ville elle-même.

Temps, sémantique, espace

Les entités géographiques sont définies non seulement dans l'espace, mais également dans le temps, et sont associées à une sémantique particulière (Plewe, 2002) :

chaque entité *existe à un endroit donné, pendant une période donnée et possède des propriétés particulières*. Ces trois éléments définissent les trois dimensions d'une entité géographique : **sa localisation et sa forme, sa période de vie et ses attributs**. Ces trois éléments forment alors l'étendue de l'entité géographique. Par exemple, la rue *Au Maire*, située dans le 3^{ème} arrondissement de Paris existe depuis 1280 (Lazare et Lazare, 1844). L'entité géographique correspondante peut être alors constituée de la forme de la rue, sa période de vie étant l'intervalle ouvert $[1280;2013[$, et sa description peut être son nom (la figure 1.5 donne un aperçu des transformations de la rue Au Maire).

Cette triple considération (sémantique, spatiale et temporelle) est présente depuis longtemps dans le domaine de la géographie (Berry, 1964; Sinton, 1978; Hägerstrand, 1970), mais les SIG se sont formés sur le modèle d'une information atemporelle (Goodchild, 2008). C'est pourtant cette information triple qu'il est crucial de manipuler pour représenter un phénomène historique, que nous désignerons dans l'ensemble de ce document comme **géohistorique** (Plewe, 2002). L'information géohistorique n'est donc pas opposée à l'information géographique, mais précise seulement le fait qu'elle traite d'entités spatiales disparues ou ayant une épaisseur temporelle donnée.

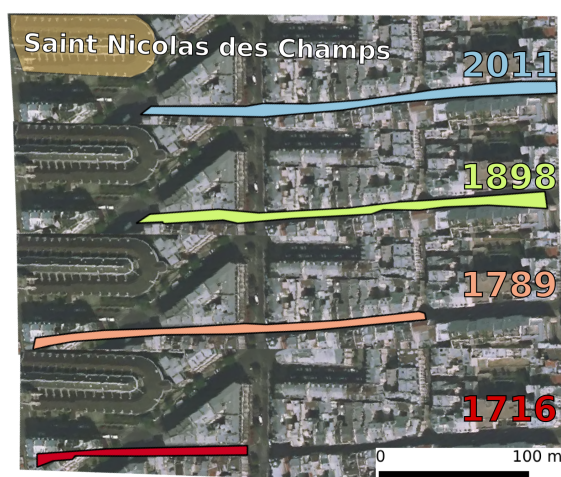


FIGURE 1.5 – Les transformations de la rue actuelle *Au Maire* comme décrite par des plans successifs de 1716 à nos jours. L'église Saint Nicolas des Champs, au sud du Conservatoire National des Arts et Métiers, sert de repère.

Identité

Une quatrième dimension doit être ajoutée à cette définition. En effet, l'identification d'un objet géographique du monde réel repose sur la possibilité de lui attribuer une existence qui le définit de façon univoque. Cette identité est définie indépendamment des trois autres dimensions. L'identification claire d'une entité est une question complexe, notamment pour les données spatio-temporelles (Del Mondo, 2011). Si l'on reprend le cas de la rue *Au Maire*, la définition de son identité n'est

pas triviale. En effet, la rue a été considérablement allongée en 1833¹⁰, englobant le tracé d'un ancien *Cul de Sac de Rome*. Dès lors, la rue *Au Maire* avant allongement est-elle la même entité qu'après ? Car si le nom reste inchangé, le tracé n'est plus le même.

L'affectation d'une identité à des entités géographiques pose d'importants problèmes lorsque l'on manipule des objets spatio-temporels. En effet, si la rue *Au Maire* conserve son identité malgré les travaux, il s'agit toujours du même objet qui se transforme. Mais si l'on considère que par cette opération c'est une nouvelle rue qui est créée, alors la question est toute différente, puisqu'il s'agit cette fois d'une entité qui disparaît au profit d'une autre, il n'y a plus transformation mais brisure. Puisqu'il s'agit dans cette thèse de traiter d'entités géohistoriques et donc spatio-temporelles, cette question se verra approfondie dans les chapitres 5 et 6.

1.2.2 Typologie des sources d'information historiques sur l'espace urbain

Puisqu'il s'agit d'espaces disparus, il n'est plus possible d'en obtenir une représentation par l'observation directe et celle-ci doit être reconstituée à partir des informations géohistoriques qui subsistent au travers de sources différentes, documents créés par les contemporains ou traces archéologiques¹¹. Ainsi, une carte topographique ancienne, un recensement comprenant les adresses postales des individus, ou les bornes de délimitation d'une ancienne censive constituent autant de ces ressources que l'on nomme ici **sources géohistoriques**. Si elles sont le matériel essentiel d'une étude de la ville, elles constituent un ensemble fortement hétérogène en raison de leur différences, mais également de biais issus des partis pris et objectifs de leurs auteurs (Knowles et Hillier, 2008).

Nous proposons ici de mettre en place une typologie de ces sources géohistoriques afin d'en expliciter les spécificités en termes de modes de représentation de l'espace mais également de biais. Il n'existe pas de classification unique de ces sources, mais il est possible de produire une synthèse des propositions apparaissant dans divers ouvrages. Nous nous fondons ici sur deux typologies en particulier qui présentent l'avantage de se compléter.

D'abord, on peut effectuer une première distinction à partir de celle établie par Graff et Barzun (1985), formant deux catégories de sources :

- **Enregistrements** : Il s'agit des sources qui fournissent une information spatiale intentionnellement, de façon plus ou moins explicite. On retrouve dans cette catégorie l'ensemble des documents d'archives portant des informations spatialisées.
- **Reliques** : Il s'agit des sources fournissant une information sur un espace passé de façon non intentionnelle. Il s'agit ici des artefacts archéologiques,

10. Index des rues fourni par la ville de Paris.

11. Une source ou document historique est [...] toute trace, tout indice de la présence, de l'activité, de la mentalité de l'Homme d'autrefois (Godding, 1975)

mais également des formes urbaines persistantes ou des éléments architecturaux de l'espace présent.

Ensuite, Arnaud (2008) tout comme Laurent *et al.* (1979) effectuent une distinction au sein des sources d'archives entre celles qui fournissent une information spatiale directement représentée et celles dont la localisation doit être retrouvée (par exemple, les adresses d'individus). Le schéma 1.6 organise ainsi les différents types de sources géohistoriques permettant de décrire l'espace urbain ancien, chaque type étant décrit plus en détail de façon à souligner ses complexités intrinsèques.

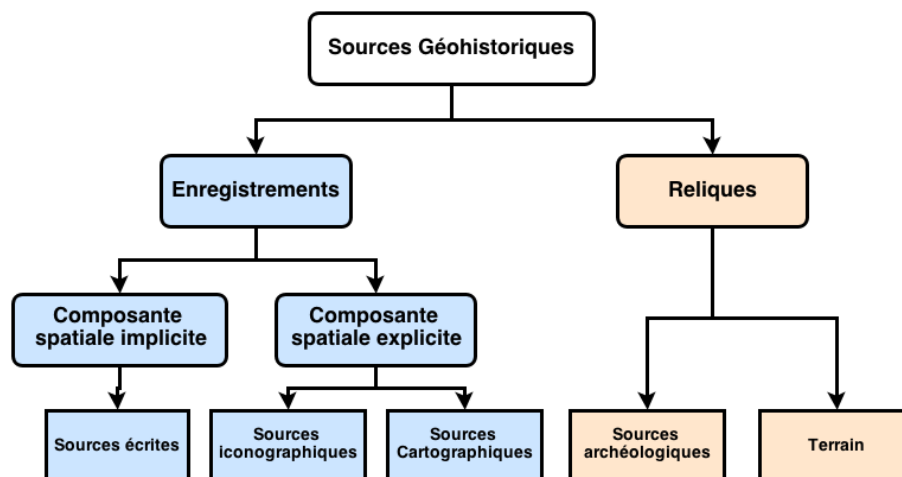


FIGURE 1.6 – Typologie des sources d'informations géohistoriques sur l'espace urbain ancien

Les sources cartographiques

Lorsque l'on se concentre sur la ville, les sources cartographiques anciennes deviennent les sources d'information principales car elles permettent de localiser l'information issue d'autres sources sans composante spatiale explicite. Par exemple, un plan parcellaire portant les numéros des immeubles permet de localiser des individus dont l'adresse est indiquée dans une source textuelle (un bottin ou un recensement par exemple).

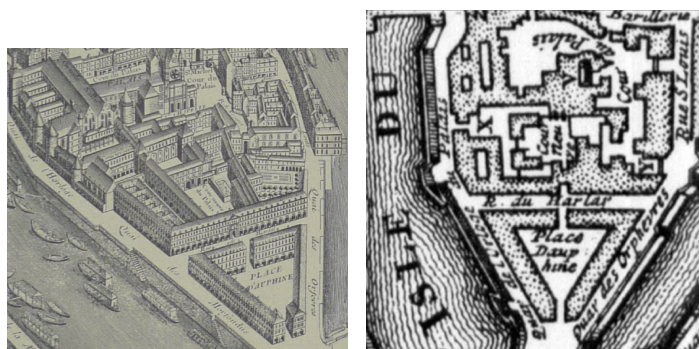
Une source cartographique est qualifiée d'historique dès lors qu'elle est située suffisamment loin dans le passé pour représenter des transformations significatives pour une étude donnée. Il s'agit donc d'une qualification subjective liée au contexte de chaque étude (Leyk *et al.*, 2005).

Les représentations de l'espace urbain varient fortement et dépendent de l'évolution des visions de la ville par les auteurs des cartes (Arnaud, 2008). À partir de la période moderne, deux types de représentations se succèdent. Le premier type correspond aux plans en perspectives, mêlant une structure générale plane enrichie de vues en élévation des structures urbaines où la précision géométrique ne constitue pas le but principal. À partir du XVIII^e siècle le plan en perspective laisse place à

des plans géométriques visant à décrire la topographie de la ville et aboutissant à des vues abstraites de celle-ci.

Ces *plans de villes* fournissent une vue globale du tissu urbain ancien, souvent matérialisée par la voirie, le parcellaire, le bâti et les espaces privés vides (cours, jardins, champs). Ce sont également des sources riches en toponymes servant de support à la localisation d'informations issues d'autres sources. Sans dresser une liste exhaustive, on retrouve souvent dans les plans de villes les noms de rues et de places, de bâtiments publics ou religieux ou encore de quartiers. La figure 1.7 illustre ainsi le changement entre les deux types de représentations au travers de deux plans de Paris datant de la première moitié du XVIII^e siècle.

Nous avons montré dans les sections précédentes que l'utilisation de l'espace urbain en histoire passe généralement par l'analyse visuelle et la production de cartes sur l'espace ancien. L'importance du rôle joué par les sources cartographiques anciennes apparaît immédiatement, car elles constituent souvent des représentations graphiques globales de l'espace étudié. Ce rôle sera précisé dans la section 1.2.3.



(a) Plan de Turgot (Bretez, 1739), 1739

(b) Plan de De L'isle (De L'isle, 1716), 1716

FIGURE 1.7 – Deux visions différentes du même espace urbain. Extraits de plans de Paris sur l'île de la cité. Les dates indiquées correspondent à la date de première publication connue.

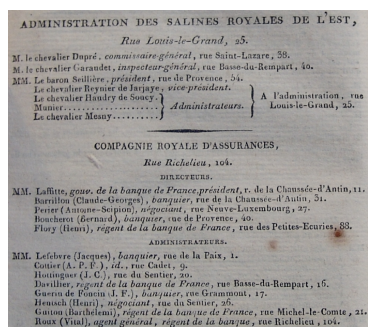
Documents écrits

Les documents écrits constituent la ressource historique principale pour les temps les plus récents¹². Le niveau de structuration de l'information géohistorique varie largement au sein de cette catégorie. Ainsi, il peut s'agir de textes littéraires, juridiques, administratifs ou bien de descriptions de lieux ou de récits de voyages où l'information géohistorique doit être interprétée par l'historien afin d'être représentable cartographiquement. Par exemple, Øyvind (2013) développe une méthodologie visant à cartographier des lieux en Norvège décrits de façon vague au sein d'un recueil de textes. Les lieux, identifiés par leurs noms, y sont généralement positionnés

12. Pour Paris, les documents écrits deviennent rares dès la période du moyen âge, les historiens devant alors avoir plus largement recours aux données archéologiques

uniquement de façon relative par leurs orientations et leurs distances par rapport à d'autres lieux.

On retrouve également des documents dans lesquelles l'information géohistorique est plus structurée et ainsi plus directement identifiable : il s'agit de listes, annuaires, tableaux, inventaires, etc. Ainsi, les bottins professionnels de Paris (dont un extrait apparaît dans la figure 1.8a) fournissent le nom, mais surtout l'adresse des travailleurs¹³, facilitant ainsi la localisation des lieux de travail. De même, un inventaire après décès peut aider à localiser les biens immobiliers d'un individu. Si ces documents sont fondamentaux, l'utilisation de l'information géohistorique qu'ils contiennent n'est pas aisée. En effet, elle n'est pas référencée directement et doit donc être positionnée dans l'espace par l'historien (Arnaud, 2008).



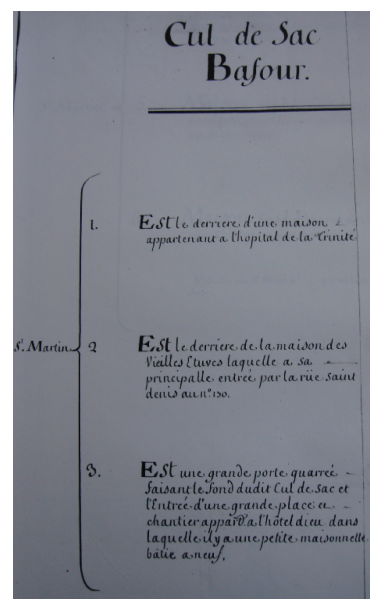
(a) Bottin professionnel, 1818.

BASFOUR (PASSAGE).

Commence au passage Saint-Denis; finit à la rue Saint-Denis, n° 309 et 302. Le dernier impair est 47; le dernier pair, 44. Sa longueur est de 406 m. — 6^e arrondissement, quartier de la Porte-Saint-Denis.

Au milieu du XIV^e siècle, on le nommait *ruelle sans chef aboutissant à la Trinité*. Vers la fin du même siècle, on commença à lui donner le nom de *Basfour*, en raison d'un four situé dans ce passage. Sa largeur actuelle varie de 2 m. 60 c. à 5 m. — Éclairage au gaz (comp^e Française).

(b) Dictionnaire des rues de Paris, 1844.



(c) Terrier du Roy, 1700

FIGURE 1.8 – Extraits de sources écrites porteuses d'informations spatiales. Source : La-DéHis, EHESS.

Iconographies et photographies

Les iconographies et photographies de la ville jouent un rôle fondamental dans la perception de l'espace urbain ancien car si elle décrivent généralement un espace restreint, elles se placent généralement du point de vue du piéton pour fournir un témoignage visuel unique de l'architecture et de l'organisation de la ville. Les sources iconographiques forment un ensemble hétérogène de documents, en raison de l'évolution de techniques mais aussi des objectifs variés des auteurs. Ainsi, les esquisses, gravures ou estampes (voir figure 1.9a) fournissent à la fois des descriptions détaillées de l'espace dépeint, mais forment aussi un puissant vecteur de messages politiques. La quantité et la qualité des estampes augmentent par exemple largement

13. L'adresse constitue un système de référence spatial indirect (Denègre et Salgé, 1996)

sous la Révolution Française, reflétant ainsi leur utilisation comme outil politique à un moment où la notion d'opinion publique prend de l'importance (Bosséno, 1990). Les plans architecturaux constituent un autre ensemble où la priorité réside dans la description précise d'une structure bâtie.

À partir du milieu du XIX^e siècle, l'arrivée de la photographie apporte un témoignage plus direct des structures urbaines (voir figure 1.9b). Cependant la photographie n'est pas un objet neutre, le message véhiculé par l'auteur se transcrivant dans les choix de cadrage, les modes de prise de vue, etc.¹⁴

Si la richesse des ressources iconographiques permet de produire une représentation détaillée de l'espace urbain, il s'agit d'une représentation partielle et dont il est souvent difficile d'évaluer la fiabilité.



(a) Le carroussel du Louvre, Paris, 1849



(b) La rue de l'arbalète, photographie de Charles Marville, entre 1865 et 1867

FIGURE 1.9 – Sources iconographiques

Sources archéologiques

Les sources archéologiques désignent l'ensemble des artefacts des structures urbaines anciennes issus de travaux de fouille, permettant une restitution physique de la ville non plus au travers de documents dont les biais peuvent être difficilement identifiables, mais au travers de données archéologiques restituant plus objectivement les bâtiments anciens sur le parcellaire actuel. Il s'agit aussi d'objets complexes, fragmentés, accumulés au cours du temps, permettant surtout une étude locale plutôt qu'un raisonnement à l'échelle de la ville (Galinié, 2000). Différents travaux portent sur l'utilisation de tels artefacts pour reconstituer un espace passé, parmi eux les travaux de Rodier et Saligny (2010), De Runz (2008) ou Stefani (2010).

Le terrain de la ville actuelle

L'espace de la ville actuelle est le produit de la sédimentation au cours du temps d'une multitude de tissus et de structures, il est un *conservatoire temporel* (Lepetit,

14. On peut citer pour illustration le travail effectué par Sicard (1999), notamment sur les photographies effectuées pour le compte du baron Haussmann par Charles Marville (1813-1879).

1996). La ville actuelle comporte alors les traces de tous ces espaces passés, directement accessibles lorsqu'il s'agit d'objets urbains ou d'éléments architecturaux ayant survécu à l'épreuve du temps, ou bien indirectement à travers la persistance dans le tissu urbain de structures anciennes (Ducom, 2009). La ville actuelle constitue pour l'historien une source d'informations à part entière, mais aussi un appui pour comprendre les structures urbaines décrites dans des sources anciennes.

1.2.3 Le plan de ville au cœur de la représentation visuelle d'un espace urbain ancien

1.2.3.1 Le socle d'une représentation de l'espace ancien

Processus de communication cartographique

Nous avons abordé en début de section la notion de représentation cartographique, dont nous avons présenté succinctement l'adaptation à un travail historique qui se fonde sur les traces survivantes d'un état ancien de la ville. Les sources cartographiques sont des traces particulières en ceci qu'elles sont des représentations visuelles concrètes et globales des anciennes structures, à la différence par exemple d'un texte qui, aussi précis qu'il soit dans sa description des lieux, fait appel à l'imagination du lecteur.

Une carte ancienne est également un objet complexe. En effet, elle est une trace et par conséquent il existe une rupture temporelle entre le moment de sa création et le moment de sa lecture (Krämer, 2010). Le créateur et le lecteur de la carte sont différents et appartiennent à des époques également différentes. Qu'il s'agisse de construire une carte ou de la lire, les deux opérations font appel à des processus cognitifs de perception de l'espace qui, par plusieurs phases de filtrages (des sens, culturels, etc.) déforment la réalité. Sans aller plus loin sur ces processus (pour lesquels on pourra se référer à Bailly (1977, 1985)), il faut prendre en considération le double filtrage qui résulte : de la représentation cartographique d'un espace par un cartographe ; de sa lecture plus tardive par une autre personne dans un contexte culturel et historique différent.

Ce processus de transmission d'un message et les déformations qu'il subit lors de son codage au sein d'une carte puis son décodage, sont abondamment abordés dans le domaine de la cartographie (Kitchin *et al.*, 2013). Afin de clarifier notre propos, nous proposons de nous appuyer sur un modèle de communication cartographique proposé par Salitchev (1978), résumé dans le schéma 1.10. Dans ce modèle, une part de la réalité est d'abord observée par le cartographe qui souhaite la cartographier (1). Il effectue alors une sélection des éléments qu'il souhaite représenter et réunit les informations nécessaires à leur description par divers moyens d'acquisition (observation visuelle, levé topographique, etc.)(2). L'information géographique alors constituée est encodée et éventuellement généralisée au sein d'une carte au moyen de symboles cartographiques (3). La carte, une fois construite, peut être décodée (lue) par un individu qui possède une connaissance des symboles cartographiques utilisés, lui permettant d'obtenir une somme d'informations géographiques. Finalement, le lecteur interprète ces informations en fonction de son expérience et de l'état de

ses connaissances sur la réalité représentée, agrandissant alors ses connaissances sur cette réalité (4). Cependant, puisque les connaissances et l'expérience varient d'un lecteur à l'autre, la part de réalité ainsi enrichie varie également.

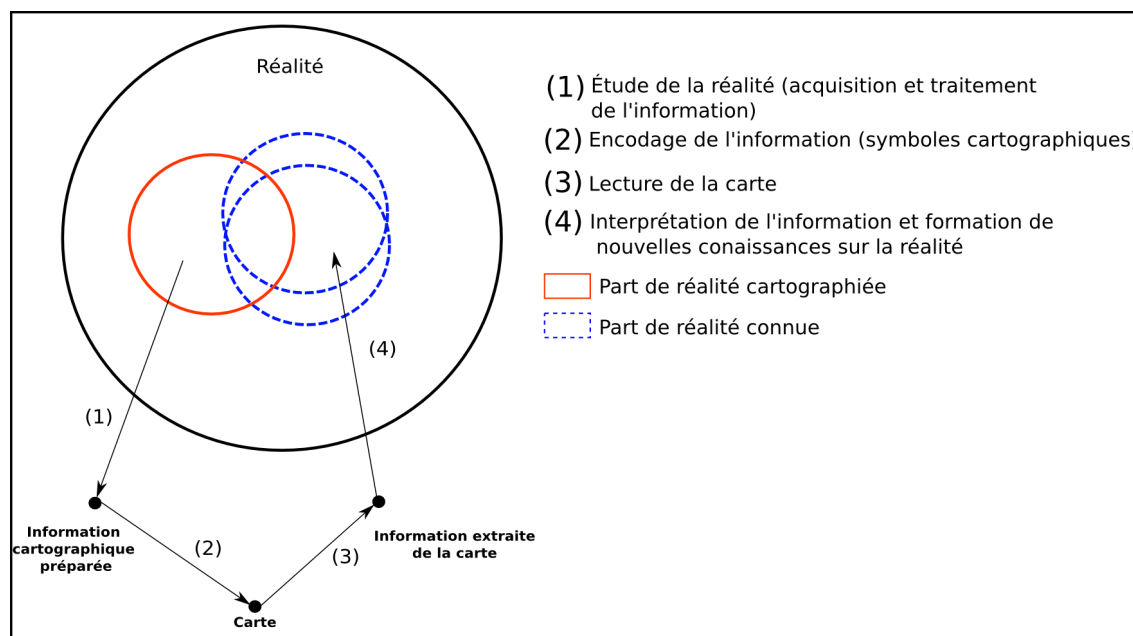


FIGURE 1.10 – Le processus de communication cartographique de Salitchev (1978)

Processus de communication cartographique différée

Lorsqu'il s'agit de lire une carte ancienne, la rupture temporelle existant entre la production de la carte et sa lecture crée deux spécificités. Nous proposons alors d'adapter le schéma de Salitchev (1978) dans le cadre d'une lecture de carte ancienne en modélisant une communication différée. Bien sûr, comme nous l'avons abordé précédemment, toute communication cartographique génère un décalage entre le message transmis par le cartographe et sa réception par un lecteur. Cependant dans le cas de cartes historiques le décalage est tel que la réalité du cartographe et celle du lecteur sont sensiblement différentes. Cette rupture amène deux particularités :

- Le **temps écoulé** entre la production de la carte et sa lecture est tel que cette dernière ne peut arriver qu'après une phase de préservation de la carte. Il peut s'agir simplement d'un archivage du document cartographique, mais aussi d'une survivance du message au travers de réutilisations plus tardives. (Pinon *et al.*, 2004, p. 10) en fournit un exemple parlant : 'original du premier plan de Paris dressé vers 1523-1530) a été perdu, mais il reste connu au travers de diverses copies.
- L'augmentation de la connaissance ne porte pas sur la réalité du lecteur, mais sur une **réalité antérieure** à laquelle il n'a pas accès. Dès lors, il dispose d'une base de connaissances plus faible pour interpréter le message cartographique transmis. Le risque est alors d'accorder un crédit trop grand, par manque d'informations, à une carte ancienne dont le degré de réalisme dépend fortement des biais de représentation (volontaires ou non) introduits par le cartographe.

Le schéma adapté est présenté par la figure 1.11. La réalité n'est plus représentée comme figée dans le temps, mais le cylindre indique son évolution. Le déroulement de la communication cartographique se déroule alors également dans le temps, la hauteur de chaque point et flèche indiquant ce déroulement. L'étape (3) décrit la phase de préservation du document, tandis que la flèche qui revient dans le temps (5) illustre l'acquisition d'une connaissance non pas sur la réalité présente du lecteur, mais sur celle passée du cartographe. Finalement, le dégradé en (1) illustre la phase d'acquisition de l'information géographique, qui elle aussi n'est plus instantanée mais dure pendant un certain temps : l'information encodée est alors non plus un instantané de la réalité mais une accumulation de temporalités différentes. Les portions de réalités perçues ou élargies ne s'étalent alors pas seulement dans l'espace, mais également dans le temps.

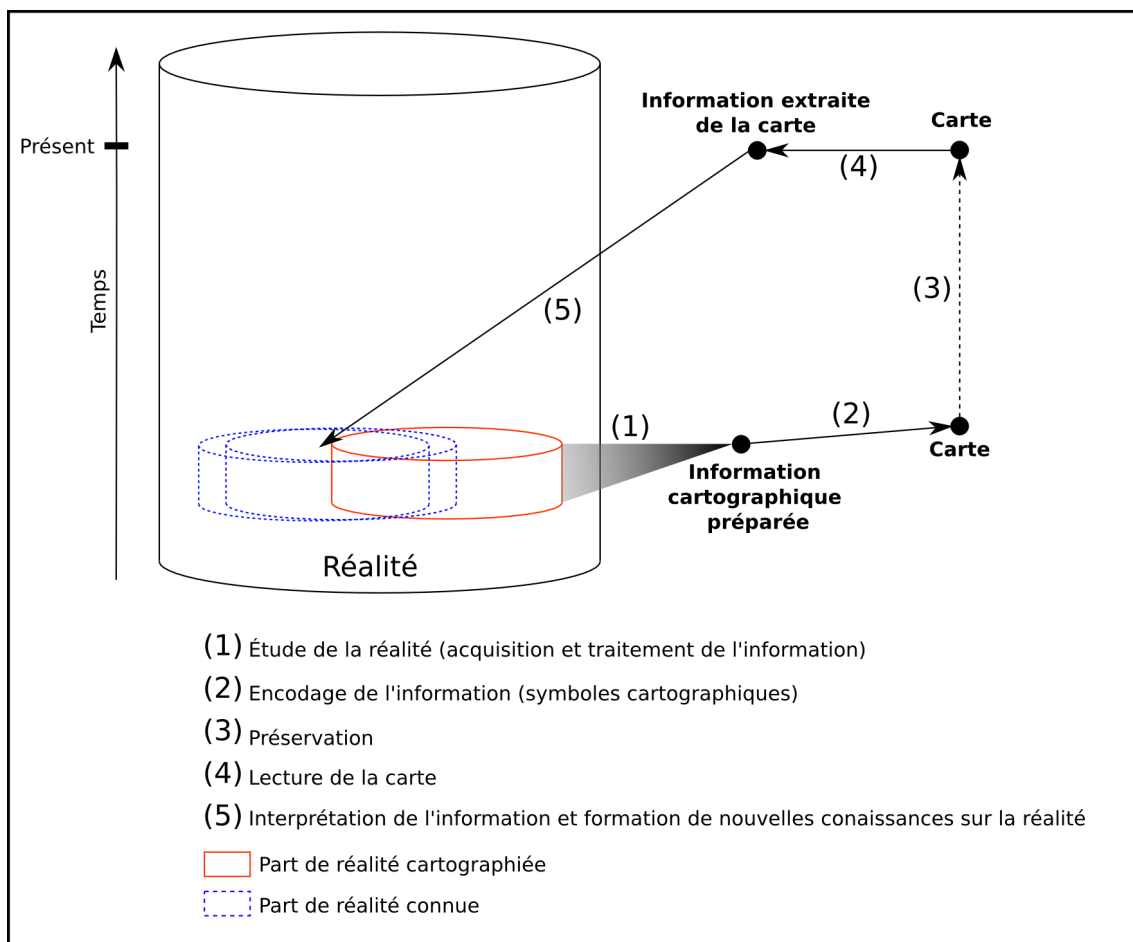


FIGURE 1.11 – Le processus de communication cartographique de Salitchev (1978) adapté pour la lecture de traces historiques dans des cartes.

1.2.3.2 Types de plans

Les sources anciennes sont des *miroirs déformants du passé* (Spitzer, 2000). Ces déformations résultent de diverses causes que nous aborderons en détails dans la sec-

tion 1.4. Cependant, les effets du temps dans la conservation du message cartographique ou dans sa construction même sont des facteurs de déformations important. Outre le temps, la multitude d'objectifs définis par les commanditaires des plans ou les cartographes eux-même peuvent impacter fortement le la précision d'une représentation cartographique. Afin d'identifier les principales différences entre plans de villes et par là les causes de déformations du passé, nous nous appuyons sur une séparation en deux classe basée sur les objectifs des plans, proposée par Pinon *et al.* (2004).

Plans fonciers

Les plans fonciers désignent l'ensemble des plans, sur l'ensemble de la ville ou un espace restreint, dont l'objectif est de décrire la répartition spatiale des propriétés foncières, c'est à dire le tissu parcellaire de la ville. À Paris, les plans parcellaires apparaissent à partir du XVII^e siècle, permettant une représentation plus claire des limites de propriétés jusqu'alors matérialisées par des repères divers telles des bornes. Arnaud (2008) distingue les plans parcellaires selon trois contextes d'utilisation : juridique, fiscal et administratif.

Dans le contexte juridique, les plans parcellaires visent à aider la résolution de conflits entre deux parties revendiquant, par exemple, des droits sur un même emplacement¹⁵. Les plans parcellaires fiscaux servent quant à eux d'outil pour la répartition et le calcul d'un impôt. Ainsi dans le Paris de l'ancien régime apparaissent à partir du XVII^e siècle les plans des censives, dressés par différentes seigneurs fonciers laïcs ou ecclésiastiques ayant droit de perception du cens sur leurs possessions parisiennes¹⁶.

On retrouve également dans cette catégorie le Papier-terrier du Roi, exécuté de 1700 à 1726-1727, décrivant pour Paris et ses faubourgs chacune des parcelles quelque soit le propriétaire foncier la possédant associé à un plan détaillé de chaque zone. Enfin, il s'agit aussi de plans administratifs précis, servant alors de base à des travaux cartographiques. Si son but est également fiscal, rentre dans cette catégorie le cadastre napoléonien dont la justesse géométrique et l'échelle (1/2000 sur Paris) en fait une base pour la construction de plans de la ville.

Plans topographiques et projets

Les plans topographiques sont des outils d'aménagement de la ville, destinés à faciliter la compréhension et la mise en place d'interventions urbaines. Ils sont une représentation épurée de la ville, l'accent étant placé sur la justesse géométrique et la fidélité à la morphologie des éléments urbains. Les plans projets visent la même précision géométrique, mais leur but diffère. Ces plans sont directement liés à un projet urbain, préparant, présentant ou faisant le bilan du projet.

15. C'est par exemple le cas du plan de Quesnel à Paris

16. Pour plus d'informations sur les censives parisiennes, voir les travaux de Weiss (2009)

Si la distinction est faite entre plans topographiques et plans projets, il n'est pas évident de distinguer les deux, les plans pouvant mêler différents messages pour livrer une vision d'un espace urbain certes précise, mais modifiée par des projets liés aux commanditaires du plan (Pinon *et al.*, 2004). Par exemple, le plan de Verniquet vise et atteint une grande précision géométrique en fournissant une description riche et complète de Paris. Mais il aussi porte le tracé de différents projets de percements de voies qui ne seront pas toujours réalisés.

1.2.3.3 Complexités temporelles des plans anciens

Le temps du plan

À la différence d'autres représentations visuelles comme la photographie, la carte ne représente pas la configuration de l'espace à un moment donné, bien qu'elle en offre une représentation statique. En effet, la carte prend nécessairement un temps variable pour être construite, pendant lequel l'espace qu'elle représente continue à se transformer. Si cet aspect peut éventuellement être négligé lorsqu'il s'agit de cartes récentes, la durée de levé des plans anciens, s'étalant souvent sur plusieurs années, oblige à prendre en compte cette temporalité interne au plan. La source cartographique ancienne doit alors être considérée comme un empilement de fragments de l'espace représenté à des moments différents. À un plan ne correspond donc pas un instant, mais plutôt une période pendant laquelle le plan est considéré comme une représentation fiable de l'espace.

Cette période ne peut être définie qu'au cas par cas, mais nous considérons par défaut qu'il s'agit de la période pendant laquelle le plan est levé, correspondant au moment où les mesures trigonométriques sont effectuées sur le terrain, et donc où le décalage temporel entre le plan et l'espace réel est moindre. Qu'elle corresponde au levé du plan ou soit définie par un exemple, cette durée correspond au *temps valide* des bases de données temporelles (Nguyen et Meier, 2006, p. 189) : *un instant précis auquel se produit un événement, [ou] un intervalle de temps qui définit la durée de validité des valeurs des données associées*. Nous nommons donc cette période **temps valide du plan**.

Types de plans par époque

Au delà de cette spécificité, Pinon *et al.* (2004) distinguent deux catégories de plans anciens selon leurs modes de création mais également selon leurs époques :

- **Plans thématiques** : Il s'agit de plans commandés par un individu ou une institution, et confiés à un cartographe. Répondant aux demandes des commanditaires, ces plans sont orientés vers une représentation mettant en valeur un élément particulier de la ville ou la cité dans son ensemble, sans forcément de considération profonde pour la justesse géométrique des tracés. À Paris, les plans de Gomboust, De La Caille et Bullet en constituent des exemples pour les XVII^e et XVIII^e siècles, ceux de Vasserot et Bellanger ou de Alphand pour le XIX^e siècle.
- **Plans topographiques** : Il s'agit souvent de plans dont les levés sont prévus par le cartographe lui-même, lequel cherche ensuite un client prêt à financer

la production de la carte. Ces plans visent à plus d'exactitude géométrique et représentent la ville de façon plus neutre, s'attachant d'avantage aux structures même de l'espace qu'à d'éventuels enjeux politiques. Dans cette catégorie rentrent notamment des plans tels que ceux de Delagrive, Verniquet ou encore Delisle, tous intégralement ou en partie financés sur fonds propres du cartographe.

Bien sur, les frontières entre ces deux catégories ne sont pas étanches. Par exemple, le plan de Jacoubet (Jacoubet, 1839) est une commande du préfet de la Seine visant la précision géométrique et une description relativement épurée de la ville. De plus, cette classification se base sur une distinction entre commanditaires "officiels" et initiatives privés. Cependant, les commanditaires de nombreux plans sont inconnus, tels que ceux de Benedict Vassalieu (De Vassalieu, 1609) et Albert Jouvain de Rochefort (De Rochefort, 1674) sur Paris. D'après Pinon *et al.* (2004), les dédicaces indiquées dans ces plans ne constituent en rien une désignation suffisante d'un éventuel commanditaire. Enfin, certains plans sont à mi-chemin des deux catégories, comme le plan de Verniquet (Verniquet, 1799), commencé par une initiative privée puis financé par les pouvoirs officiels. Malgré la porosité de cette frontière, la considération de ces différents éléments nous permet de préciser les complexités engendrées par la temporalité propre aux plans.

Multiplicité temporelle

Qu'il s'agisse de plans thématiques ou purement topographiques, il peut coexister au sein d'un plan des représentations issues de périodes différentes présentées sur un même fond de carte. La figure 1.12 illustre un tel cas sur différents plans thématiques présentant les interventions urbaines effectuées dans Paris au XIX^e siècle. Dans le cas de la carte topographique, bien que l'objectif soit de représenter la ville au moment du levé, le plan comporte également des tracés de rue en projet, dont le percement n'est pas commencé et qui n'ont donc aucune réalité physique au moment du levé du plan. Pour la carte thématique, il s'agit cette fois de représenter un historique et des événements antérieurs sur un plan de Paris, résultant ici aussi en une accumulation de temporalités au sein de la source.



FIGURE 1.12 – Différentes temporalités au sein d'une même carte : les travaux de voirie effectués et édifices construits de 1789 à 1889, d'après Huet (1889)

Durée des opérations de levés

Comme nous l'avons introduit, le tracé d'un plan de ville complet n'est pas une opération rapide et nécessite une ou plusieurs campagnes de levés, étalant la période de construction du plan sur plusieurs années alors que l'espace urbain continue de se transformer. Lors de la production finale du plan, celui-ci contient alors des informations provenant de phases d'observation de la ville à des moments différents, pouvant mener à des anachronismes. Par exemple, le plan de Verniquet (Verniquet, 1799) fit d'abord l'objet d'un levé à partir de 1775, suivi par un second entre 1785 et 1795.

Réutilisation de plans

Lever un plan étant une opération longue et coûteuse, la construction d'un plan est parfois effectuée sur la base d'un plan plus ancien. Sur Paris par exemple, de nombreux plans du XIX^e siècle réutilisent le plan de Verniquet tracé au moment de la Révolution française, ce qui est notamment le cas de plans très largement utilisés par l'administration de l'époque, comme l'Atlas de Jacoubet (Jacoubet, 1839) ou plan de Girard (Girard et Verniquet, 1819). Cette réutilisation peut avoir comme conséquence un effet d'anachronisme qui se produit entre la période de publication du plan et l'espace qu'il représente réellement. Cette pratique induit aussi la reproduction des éventuelles erreurs présentes dans les opérations de lever et de gravure du premier plan.

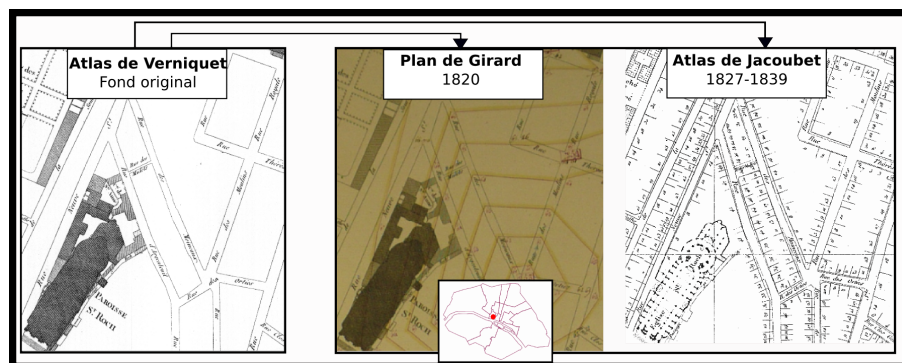


FIGURE 1.13 – Réutilisation du plan de Verniquet (à gauche) dans deux plans plus tardifs (à droite)

1.3 L'apport des SIG pour étudier l'histoire : un choix méthodologique

1.3.1 Introduction : définitions préalables

Des Systèmes d'Informations Géographiques...

Il n'existe pas une définition unique des Systèmes d'Informations Géographiques (SIG), mais différentes approches issues des diverses disciplines qui se sont intéressées aux données géographiques (Denègre et Salgé, 1996).¹⁷ La relative jeunesse de la

17. Notons dès à présent que dans les travaux présentés au sein de cette thèse, l'acronyme anglais "GIS"

géomatique ainsi que le fort attrait pour la représentation d'une information spatiale ou spatialisée au sein de divers domaines (scientifiques, industriels, éducatifs, etc.) ont engendré une multiplicité d'usages, d'applications rassemblés sous le même sigle (Chrisman, 1999).

On peut avoir une appréciation indirecte de la variété de dimensions qui marque l'horizon de ce domaine en évoquant (Denègre et Salgé, 1996) deux définitions qui synthétisent les principales approches existantes : d'une part, une vision technique des SIG issue principalement de l'informatique, d'autre part un point de vue centré sur l'utilisation de ces systèmes pour répondre à des besoins thématiques.

- La première définition est elle même une synthèse de diverses propositions issues de la géomatique, de l'informatique ou de la géographie, et qui peut être résumée par la définition adoptée par la Société Française de Télédétection¹⁸ : *un système informatique permettant, à partir de diverses sources, de rassembler et d'organiser, de gérer, d'analyser et de combiner, d'élaborer et de présenter des informations localisées géographiquement, contribuant notamment à la gestion de l'espace*. Dans cette optique, on considère également des systèmes d'informations au sens plus général du terme et pouvant être constitués de procédures manuelles. Les outils logiciels, les processus de traitement de l'information spatiale, mais également les bases de données géographiques entrent dans cette catégorie.
- La seconde définition est plus proche de l'usage fait de l'information géographique dans des domaines autres que la géomatique et se focalise d'avantage sur les données et leur exploitation. Un SIG est ainsi défini comme *un ensemble de données repérées dans l'espace, structuré de façon à pouvoir en extraire commodément des synthèses utiles à la décision* (Didier, 1990). Il ne s'agit plus de s'intéresser aux systèmes de traitement de l'information géographique eux mêmes, mais aux données structurées qui en sont issues et qui constituent la base d'un travail thématique (production de cartes, analyse spatiale, etc.).

Ces deux définitions sont particulièrement intéressantes dans la mesure où elles montrent finalement deux visions différentes des systèmes d'informations géographiques, l'une provenant des domaines à l'origine des systèmes eux même, la seconde correspondant à la vision des utilisateurs finaux de ces outils. Pourtant, elles sont aussi complémentaires, chacune se positionnant à des endroits différents du même système de manipulation d'informations géographiques. Si cette multiplicité peut sembler uniquement formelle, elle est en fait source d'une ambiguïté importante dans la désignation des SIG destinés à des travaux historiques.

... aux Systèmes d'Informations Géographiques Historiques

Les premières utilisations des SIG en histoire au début des années 1990 ont donné lieu à une explosion du nombre de travaux. C'est pourquoi on peut plutôt parler d'un ensemble d'éléments rassemblés sous une dénomination polysémique

correspond systématiquement à *Geographic Information System*. Lorsqu'un homonyme est utilisé (comme *Geographic Information Science*), l'acronyme ne sera pas utilisé pour éviter les confusions.

18. <http://www.sfpt.fr/>

que d'un élément unique. C'est pourquoi il nous paraît important de clarifier cette définition pour pouvoir nous positionner efficacement par rapport aux propositions existantes. Le plus simple est sans doute d'éclaircir cette notion à partir des énoncés que l'on retrouve dans la littérature. À ce niveau (Knowles et Hillier, 2008), les SIG historiques peuvent désigner en même temps :

- Les outils logiciels qui permettent de visualiser, stocker et analyser une information géohistorique issue de sources multiples.
- Les méthodologies d'intégration d'une information géohistorique au sein d'un outil SIG (logiciel ou base de données).

Mises en rapport avec les définitions précédentes des SIG, il apparaît clairement que ces deux propositions s'insèrent en plein dans la première vision, que nous venons d'évoquer et qui exprime une approche principalement technique des SIG, car il s'agit ici de considérer des systèmes spécifiquement dédiés à l'analyse et au traitement de données historiques. Pourtant, on le voit, on parle aussi de SIG historique pour désigner un ensemble de données qui changent dans le temps (Noizet *et al.*, 2008). Dans cette optique, nous rejoignons donc davantage la seconde vision, axée non plus sur les outils informatiques eux-même mais sur les données spatiales produites à partir de la collecte de sources d'archive ou d'observations archéologiques.

Dans leurs travaux, Gregory et Healey (2007) sortent des catégories déjà identifiées en parlant cette fois non plus des outils, des méthodes ou des données, mais de la discipline même émergeant de l'utilisation des outils SIG dans le cadre d'études historiques. Cette proposition particulière est fondamentale car elle met en lumière la position des historiens comme utilisateurs principaux. En effet, il s'agit plus souvent d'utiliser des outils existant et d'adapter leur usage à un problème historique que de mettre en place des outils SIG spécifiques. Nous le constaterons plus précisément en explorant les propositions d'outils existants dans la partie 1.5. Les trois définitions présentées ici ne sont qu'un bref aperçu des écrits sur les SIG historiques, mais elles permettent de mettre en lumière deux besoins particuliers :

- **H** **Gérer des sources et des données hétérogènes**
Nous avons vu dans la section 1.2.2 que les sources historiques sont de nature, de forme, de contenu variés, induisant une forte hétérogénéité des données qui peuvent en être extraites. Nous avons vu dans la section 1.2.2 que les sources historiques, par la variété de leur forme et leur contenu, induit une forte hétérogénéité des données qui peuvent en être extraites. Par conséquent, puisque la nature même du travail historique implique de croiser un nombre important d'informations différentes, les SIG peuvent se révéler des outils extrêmement efficaces utilisant la localisation spatiale comme pivot pour intégrer différentes sources de données. Partant du constat que les sources historiques sont imparfaites et dont la fiabilité n'est jamais totale, Knowles et Hillier (2008) avance également le besoin d'être capable de questionner toute donnée sur l'origine et le degré de fiabilité de la source dont elle provient.
- **ST** **Modéliser et représenter les changements de l'espace dans le temps**
Comme le font remarquer plusieurs auteurs (Berman, 2009; Parker, 2010), il s'agit de traiter et analyser des *données qui changent dans le temps*, c'est à

dire des **données spatio-temporelles**. On retrouve ici l'importance de traiter non pas d'espaces figés, mais de ses transformations au cours du temps, qui doivent donc être gérées par un SIG spécifiquement dessiné pour les traitements historiques.

Il apparaît donc que ces deux éléments constituent des besoins essentiels auxquels un SIG dédié à l'étude de phénomènes historiques doit pouvoir répondre. S'il s'agit de besoins généraux, ils sont pourtant fondamentaux et ils questionnent particulièrement les systèmes existants. En effet, l'absence d'hégémonie d'un SIG particulier nous emmène à questionner l'adéquation des outils et des méthodes mis en place jusqu'à maintenant pour répondre à ces besoins. Certains travaux, comme l'étude du quartier des Halles effectuée par Boudon et Blécon (1977) montrent d'ailleurs que les outils informatiques ne sont pas strictement nécessaires pour travailler de manière conséquente sur la morphologie d'un espace ancien et ses dynamiques de transformations. Il nous faut donc explorer les raisons qui feraient des SIG des outils importants en histoire, et quels sont les objectifs précis auxquels ils peuvent répondre. Pour ce faire, nous proposons d'abord d'identifier les particularités du traitement de l'information géohistorique au sein d'un SIG, notamment quant à la manipulation de données hétérogènes et imparfaites (section 1.4), puis d'explicitier ces objectifs et leur adéquation avec le travail historique en section 1.3. De là, nous pourrons dégager un positionnement plus clair pour notre travail.

1.3.2 Représentation de l'information historique

Modes de représentation de l'information géographique dans un SIG

Les SIG s'appuient sur deux modes de représentation de l'information géographique (Denègre et Salgé, 1996) :

- Le mode **raster** représente un espace géographique sous la forme d'une grille régulière dont la résolution est fixe. Chaque cellule de la grille stocke une information unique, faisant d'elle la plus petite entité géographique représentable. Un raster permet de stocker des informations thématiques discrètes, comme le type d'occupation du sol, mais également des données continues (élévation, température, etc.). Les images constituent les données sous forme raster les plus répandues : la cellule de la grille correspond à un pixel de l'image. Dans le cadre de données historiques, il s'agit la plupart du temps de cartes historiques numérisées.
- Le mode **vecteur** permet de représenter les formes des objets du monde réel à l'aide de quelques primitives géométriques simples : points, lignes polygonales et polygones. Chacune de ces primitives est composée de points auxquels sont attribués des coordonnées géographiques. À chaque primitive géométrique peut être associé un ensemble d'attributs, permettant de stocker des informations sémantiques associées. Lorsqu'il s'agit de données historiques, les données vecteur sont souvent issues de la vectorisation, généralement manuelle, de sources cartographiques. La figure 1.14 illustre les deux modes de représentation de l'information géographique sur un extrait d'un plan de ville de Paris.

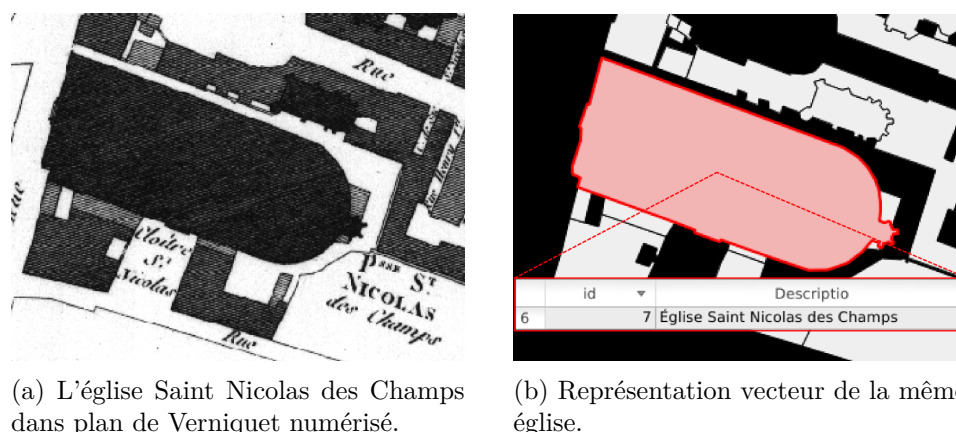


FIGURE 1.14 – Les deux modes de représentation de l'information géographique. Raster en (a), vecteur en (b).

1.3.2.1 Intégrer une information géohistorique explicite

On considère les cartes, plans et mappes, ainsi que les représentations iconographiques d'un espace ou d'un monument, comme des sources géohistoriques explicites. Les reliques, vestiges archéologiques ou traces sur le terrain, peuvent également être considérées comme explicites, leur localisation pouvant être acquise par mesure directe (en particulier par relevé GPS). Nous ne considérons pas ici le cas des sources iconographiques qui sont, aux époques considérées, généralement des vues en perspectives de bâtis, de lieux ou de scènes de vies et qui nécessitent dès lors des données 3D pour être recalées. La constitution de données 3D sur l'espace passé est problématique, en raison de la pauvreté des sources. En effet, peu de sources anciennes permettent de reconstruire des éléments en volumes et plus rares encore sont celles qui représentent ces éléments avec réalisme.

Géoréférencement

Les sources cartographiques constituent un élément central dans la construction d'une démarche qui vise à reconstituer la morphologie d'un espace passé et d'en donner une représentation globale.

Afin de pouvoir les confronter de manière critique à d'autres sources et d'autres données spatiales, il est nécessaire d'assigner une localisation précise dans l'espace géographique aux cartes et plans. Cette opération, nommée *géoréférencement* n'est pas propre aux sources historiques mais elle constitue une étape cruciale dans la constitution d'un SIG historique pour lequel une grande part des données spatiales provient de cartographies anciennes (Rumsey, 2002).

L'opération de géoréférencement implique tout simplement la tentative de faire correspondre, et donc de superposer dans un repère géographique, le plus exactement possible, les objets géographiques communs à la carte ancienne et à la carte contemporaine. Pour ce faire, la méthode habituellement adoptée (Rumsey, 2002) consiste à placer des *points de contrôle*, généralement de façon manuelle, dans la

carte numérisée sous forme raster une carte. À chaque point de contrôle est alors lié un second point situé dans un référentiel spatial muni d'un système de coordonnées de référence choisi au préalable. Les nombreuses bases de données géographiques actuellement à disposition servent souvent de référentiels, les points de contrôles étant placés sur des éléments du paysage (point culminant, tour, croisement de rues, etc.) présent à la fois dans le référentiel et dans la carte. Quant au géoréférencement, il repose sur l'application de méthodes de transformation mathématiques permettant, à partir des points de la carte ainsi associés, d'assigner une coordonnée à chaque point de la carte raster (Grosso, 2010).

Comme le note (Rumsey, 2002, p. 5), le géoréférencement parfait de cartes anciennes sur un référentiel actuel est souvent impossible. D'abord, les systèmes de coordonnées utilisés dans les cartes anciennes ne correspondent plus à ceux utilisés actuellement, et sont souvent difficiles à reconstituer. De plus, la précision géométrique des plans anciens est variable, en raison des techniques de levé ou encore de la précision des instruments de mesures géodésiques utilisés. Les biais introduits par les cartographes ayant tracé la carte influent également sur la précision géométrique des informations qu'elle contient. Ainsi, sur la carte de Cassini dressée pour l'ensemble de la France entre 1747 et 1789¹⁹, la taille et le nombre de certains symboles peuvent affecter la disposition des éléments, en provoquant, par exemple, le déplacement du tracé d'une rivière au profit du toponyme d'un village (Costes, 2012, p. 3). Enfin, les conditions de conservation peuvent déformer une carte ancienne, notamment si son support est sensible à l'humidité (bois, toile, papier, etc.).

Vectorisation

La seconde méthode d'intégration d'une information géohistorique explicite dans un SIG consiste à transformer les objets géographiques représentés dans la carte en objets vecteurs : il s'agit de la *vectorisation*. Cette étape succède au géoréférencement et permet d'assigner à chaque objet géographique identifié dans la carte une primitive géométrique, des informations sémantiques sous forme d'attributs ainsi qu'une identité. Cette identité, implicite lorsque l'objet est saisi, peut être explicitée par exemple à l'aide d'un identifiant unique. La vectorisation d'une carte ancienne offre plusieurs avantages :

- Elle permet d'extraire l'information sémantique contenue dans la carte et de s'abstraire du style cartographique qui lui est associée.
- La séparation des composantes de l'information spatiale permet de nombreux traitements automatiques par un outil SIG (croisements attributaires, requêtes spatiales, sélection, etc.)

Tout comme le géoréférencement, la vectorisation de cartes est loin d'être une tâche anodine. En premier lieu, elle implique des opérations de précision qui peuvent se révéler longues en raison du nombre d'objets, ce qui explique en partie la rareté des bases historiques vectorielles (Rumsey, 2002). En second lieu, la vectorisation d'une carte implique d'identifier des objets au sein d'une carte, c'est à dire à assigner une identité particulière à une zone de l'espace. Dans les cartes historiques

¹⁹. Les relevés sont terminés au moment de la Révolution, mais la carte ne sera entièrement éditée qu'en 1815.

cette assignation peut poser problème. En effet, dans une carte récente il peut être possible de comparer une représentation cartographique avec l'espace réel et aider ainsi la vectorisation, ce qui est impossible lorsqu'il s'agit d'une carte historique. En conséquence la carte ancienne devient le seul témoin d'un objet géographique dont l'identification n'est pas toujours aisée.

1.3.2.2 Intégrer une information géohistorique implicite

Les sources d'information géographique habituellement définies comme "implicites" sont constituées par l'ensemble des ressources portant une information spatiale qu'il n'est pas possible de localiser directement. La forme de ces informations et la difficulté de leur localisation varient largement : l'adresse postale d'un individu ou l'itinéraire d'un voyageur sont toutes deux des informations implicites et pourtant très différentes.

Géocodage

Conceptuellement proche mais techniquement différent des opérations de géoréférencement, le *géocodage* consiste à placer une donnée spatiale implicite dans un repère géographique absolu (Goldberg *et al.*, 2007) en lui attribuant une coordonnée géographique. Très utilisée dans diverses disciplines, cette opération permet :

- D'intégrer une information sémantique dans un référentiel spatial en vue de croisements avec d'autres informations. C'est notamment le cas des adresses, qui, une fois localisées, peuvent permettre de mettre en relations divers sources dont elles constituent le lien (par exemple pour des habitations parisiennes : des actes notariés, des recensements, des calepins cadastraux, etc).
- De cartographier ces informations et éventuellement de faire apparaître des relations jusqu'alors invisibles (Gauthiez et Zeller, 2010).

Les méthodes de géocodages sont très diverses et dépendent fortement du cas d'application et aboutissent au développement de méthodes *ad-hoc*, dont on pourra trouver un état de l'art détaillé dans (Goldberg *et al.*, 2007). Depuis quelques années cependant, les différentes propositions de méthodes de géocodage tendent à s'appuyer sur des index géographiques en ligne tels Geonames²⁰ ou Pleiades²¹.

1.4 Les difficultés dues aux imperfections de l'information géohistorique

Nous venons de voir que les SIG constituent un outil performant de croisement et de partage de données historiques spatiales, permettant aux chercheurs de mettre en relation des données dont l'accès et l'exploitation peuvent être difficiles. Par leur nature, les SIG nécessitent des traitements spécifiques pour intégrer et traiter des informations géohistoriques. Le géoréférencement, la vectorisation de plans et le géocodage constituent les trois outils de base permettant ces opérations. Cependant, tout en étant très performante, ces opérations sont loin d'être anodines. En intégrant une suite de techniques conçues et développées pour des traitements fondés

20. <http://www.geonames.org/>

21. <http://pleiades.stoa.org/>

sur l'observation directe des sources et des terrains, elles risquent de déformer les informations portées par les sources historiques dont la spécificité est avant tout liée à leurs imperfections. L'information géohistorique, comme toute information, est en effet toujours entachée d'imperfections (Plewe, 2002). La principale raison, déjà évoquée pour la représentation cartographique, réside dans l'impossibilité de saisir la complexité d'éléments qui permettent de saisir un phénomène dans la totalité de ses dimensions. Ces imperfections ont été largement étudiées et sont notamment décrites par Bouchon-Meunier (1995) qui en a proposé une typologie fondée sur trois classes :

- **L'incertitude** : Une information est dite incertaine lorsque on a un doute sur sa validité. Par exemple, le témoin d'un vol de voiture peut expliquer à la police qu'il "pense que la voiture était verte". L'information sur la couleur de la voiture est incertaine : elle est très probablement verte, mais ce n'est pas certain.
- **L'imprécision** : Une information imprécise aboutit à l'impossibilité de définir clairement des limites à une ou plusieurs entités. Ces limites peuvent être mal connues en raison de manques dus aux appareils de mesure (par exemple un thermomètre précis à 1°C près). Une imprécision apparaît également lorsque les limites d'une entité sont flexibles car désignées par des termes vagues. Ce peut être le cas également quand les limites sont flexibles, soit parce qu'elles sont définies de façon vague, soit parce qu'elles sont proviennent
- **L'incomplétude** : Il est possible qu'une information parvienne de façon incomplète, soit parce qu'une partie est manquante, soit parce qu'elle ne couvre pas l'intégralité du phénomène représenté.

Ces trois types d'imperfections se retrouvent à différents instants de l'acquisition, la représentation et l'analyse de l'information géohistorique. Dans cette section, nous proposons de mettre en lumière ces imperfections et de les mettre en rapport avec le processus de représentation d'un espace ancien abordé dans la section 1.2.

Les imperfections des informations spatiales représentées au sein des SIG sont aujourd'hui bien connues et identifiées au travers de plusieurs taxonomies visant à classifier les types, les causes et les effets de chaque imperfection (Fisher, 1999; Klir et Yuan, 1995; Giles M. Foody et Atkinson, 2002; Fisher *et al.*, 2005). Lorsqu'il s'agit d'informations géohistoriques, spatiales mais également temporelles, issues de sources décalées temporellement, il devient nécessaire d'adapter les approches et les modèles existants. Au cours de la dernière décennie, plusieurs chercheurs ont élaboré des propositions visant à spécialiser les modèles existants pour le traitement de données géohistoriques (Plewe, 2002; Leyk *et al.*, 2005; De Runz, 2008). En nous inscrivant à notre tour dans cette optique, à partir de l'analyse du processus d'intégration des informations géohistoriques dans un SIG, nous avons essayé de proposer un modèle adapté pour l'analyse des transformations morphologiques de l'espace urbain.

Incertaine et imperfection

Notons dès à présent que les modèles d'imperfection de l'information spatiale rassemblent les différents types d'imperfections des données et de l'information spa-

tiale sous le terme d'*incertitudes*. Cependant, de la même manière que De Runz (2008), nous lui préférons ici le terme d'*imperfection* issue de la classification de Bouchon-Meunier (1995). Le sens de l'incertitude pour Bouchon-Meunier (1995) et pour les modèles tels celui de Fisher *et al.* (2005) varient sensiblement. En effet, pour Fisher *et al.* (2005) l'incertitude est le concept qui coiffe l'ensemble des défauts qui affectent l'information, tandis la validité de l'information est plutôt liée au concept d'erreur.

1.4.1 Le processus d'intégration d'une information géohistorique

Nous avons pour l'instant parlé des imperfections qui affectent l'information géohistorique de façon générale, sans en identifier les causes, les types et les effets. La première étape est alors d'identifier les causes de ces imperfections. Dans les sections précédentes, différentes imperfections sont apparues : imperfections des sources historiques, de l'information qu'elles contiennent ou bien encore de sa représentation au sein d'un SIG. Ces divers générateurs d'informations imparfaites se trouvent répartis à différents instants de l'intégration d'informations géohistoriques dans un SIG.

Les étapes du processus

D'autres recherches ont permis d'individualiser les différents moments durant lesquels des imperfections apparaissent à l'intérieur du processus d'intégration d'informations géohistoriques dans un SIG. Ainsi, on distingue classiquement trois étapes dans ce processus (Fisher, 1999; Longley *et al.*, 2005), décrites dans la figure 1.15. La proximité avec la figure 1.11 sur la communication cartographique est naturelle : il s'agit d'une forme particulière de représentation d'informations spatiales. Plewe (2002) fournit une description plus détaillée de chacune des étapes du processus. Ainsi, d'après lui, le processus serait constitué de la chaîne d'états suivants :

- D'abord, une part de la réalité est **conceptualisée**. Cette étape consiste à identifier différentes entités géographiques qui constituent la réalité d'un lieu à une époque donnée, selon leur nature (voir le paragraphe 1.2.1). L'identification de ces entités construit leur *étendue idéale* (*ideal extent*), c'est à dire la représentation mentale des trois dimensions qui composent l'étendue de l'entité.
- Ensuite, les entités identifiées sont **mesurées** par le biais de différentes méthodes (observation, généralisation, classification, interprétation ou encodage), ce qui permet de construire leur représentation sous la forme de données manipulables par un SIG. Cette opération permet de créer l'*étendue évaluée* (*asserted extent*) de l'entité, c'est à dire la transformation de l'entité conceptuelle en donnée concrète dont les propriétés ont été mesurées.
- Enfin, la dernière phase d'**analyse** permet de construire de nouvelles entités à partir de celles déjà connues. Généralement, ces entités sont plus précisément définies que les données initiales (Plewe, 2002). Par exemple, une personne localisée par une étape de géocodage aura une étendue plus complète que la donnée d'origine dont on ne connaissait pas la localisation.

Sources des imperfections

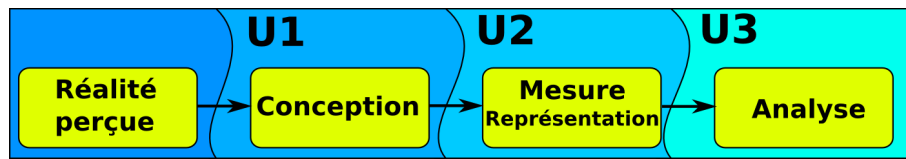


FIGURE 1.15 – Schéma des différents filtres (U1,U2,U3) déformant la réalité perçue et représentées dans un SIG. Chaque filtre génère des imperfections différentes affectant l'information spatiale transmise.

Longley *et al.* (2005) identifient trois principales causes menant à la dégradation de l'information spatiale. Leyk *et al.* (2005), en l'adaptant au cas où l'information provient de sources historiques, les présentent comme tel :

- Les **Imperfections orientées production** : Il s'agit des différentes imperfections inhérentes à l'utilisation de sources historiques. Pour Leyk *et al.* (2005), il s'agit des erreurs des sources : les erreurs de position et les erreurs dans les attributs des données (Atkinson *et al.*, 2004). Ainsi, les erreurs de levés des cartographes ou l'imprécision de leurs mesures se retrouvent invariablement au sein des sources.
- Les **Imperfections orientées transformations** : Ce sont les différentes imperfections qui résultent de la transformation de l'information contenue dans les sources en donnée utilisable, sous forme vecteur ou raster. Ainsi, le géoréférencement de cartes anciennes déforme l'espace, et la vectorisation ne permet que de saisir partiellement le contenu de la carte.
- Les **Imperfections orientées application** : Il s'agit cette fois des imperfections dues à la comparaison des données issues de différentes sources et à leur confrontation à l'espace présent. Les mauvaises interprétations du contenu des sources dues au manque d'informations sur les pratiques des cartographes en sont un exemple.

Ces différentes sources d'imperfections se retrouvent de fait sur le schéma 1.15. En effet, la transition entre les différents états de l'information géohistorique est effectuée par l'application de plusieurs filtres (dénommés U1, U2 et U3) qui s'appliquent les uns après les autres. Ces filtres déforment la réalité de façon progressive, dégradant et modifiant l'information initiale.

- Le premier filtre **U1** correspond à la perception de la réalité qui aboutit à la représentation mentale d'un espace (Bailly, 1985). Lorsque la réalité passée n'est plus perceptible que par des sources historiques, le filtre U1 génère les imperfections orientées production.
- Le second filtre **U2** correspond à la transition entre l'entité conceptuelle et l'entité mesurée. Sur des sources historiques, la mesure est effectuée par le géoréférencement, le géocodage et la vectorisation. Le filtre U2 est source d'imperfections de transformation.
- Enfin, les données géohistoriques sont interprétées au sein d'une grille d'analyse, créant un dernier filtre **U3**. Cette fois, ce sont les imperfections orientées application qui apparaissent.

1.4.2 Types d'imperfections de l'information géohistorique

En montrant le processus d'intégration de l'information géohistorique depuis des sources, nous avons mis en lumière les différents événements qui dégradent l'information. Cependant, il faut encore préciser quelles sont exactement les types d'imperfections qui peuvent affecter cette information. Dans le cas de l'information spatiale, la taxonomie de Fisher *et al.* (2005) fait école car elle associe à chaque type d'imperfection un formalisme mathématique permettant de la représenter. Sur ces bases, De Runz (2008) a proposé une taxonomie adaptée à l'information archéologique. Cette taxonomie s'adaptant à nos données, nous proposons de l'utiliser pour identifier et traiter l'imperfection de l'information géohistorique dans cette thèse. La figure 1.16 présente cette taxonomie. Nous détaillons ensuite les différents types d'imperfection en les spécialisant sur le cas des données géohistoriques issues de cartes anciennes.

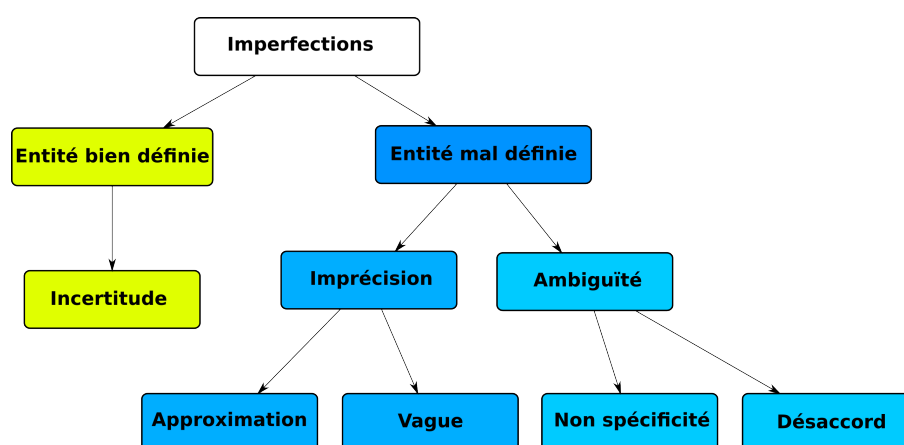


FIGURE 1.16 – Taxonomie des imperfections de l'information géohistorique, d'après De Runz (2008)

1.4.2.1 Détails de la taxonomie

Qualité de définition d'une entité

Une première distinction est faite selon la qualité de définition de l'entité géohistorique considérée. En effet, le premier problème qui se pose lors de l'étape de conceptualisation est la construction des limites de l'entité, ou, pour reprendre Fisher *et al.* (2005), de l'affectation d'une classe à chaque portion de l'espace. Une entité est bien définie lorsque qu'elle est spécifiée de manière *complète, précise et unique* (De Runz, 2008). Afin d'illustrer chaque catégorie d'imperfection, nous nous appuyerons sur l'exemple de la rue *Au Maire* illustrée dans la figure 1.5. Chaque type d'imperfection peut s'appliquer à toute une entité ou seulement à l'une de ses trois composantes.

Incertitude

Même lorsqu'une entité est bien définie, il est possible qu'elle soit incertaine. Pour se rattacher au sens fixé par Bouchon-Meunier (1995) (1.4), il peut s'agir d'une entité dont la validité d'au moins une partie de son étendue est mise en doute.

Reprenons le cas du champ de blé dont l'information sur l'emprise spatiale et le type de culture est issue d'un plan ancien. Son étendue évaluée correspond donc spatialement à son emprise sur la carte, sa période de vie à la date du plan et sa sémantique est donnée par son type de culture. Cette dernière information peut être soit fausse, soit vraie. Elle peut être affectée par diverses erreurs et la source en elle-même peu fiable (De Runz, 2008, p. 28). Dès lors, la validité du type de culture *blé* peut être mise en doute, bien que le champ soit une entité bien définie.

Imprécision

Une entité peut être mal définie lorsque la spécification de ses éléments constitutifs n'est que partielle. Plus précisément, l'imprécision apparaît dès lors qu'il est impossible de fixer des limites bien définies pour une entité ou pour un de ses éléments, rendant impossible sa catégorisation simple dans une classe unique et menant à la nécessité de classer les limites comme floues. Reprenant Bouchon-Meunier (1995), citeDeRunz2008 spécialise l'imprécision en deux sous-catégories : le vague et l'approximatif.

- Le fait qu'une entité soit définie de façon vague résulte de l'utilisation de termes du langage naturel subjectifs ou flexibles (De Runz, 2008, p. 29). Par exemple, on peut dire que l'église Saint Nicolas des champs est "proche de" ou bien même "à l'entrée ouest" de la rue *Au Maire* et "au nord" de celle-ci. Ces trois termes permettent de localiser l'église par rapport à la rue de façon indirecte, mais uniquement de façon vague. Cette description ne permet pourtant pas de localiser de manière précise l'église car chacun des termes ne définit que des relations spatiales variables et subjectives.
- Une entité est définie de façon approximative lorsque ses limites ne peuvent pas être clairement fixées, en conséquence de quoi l'entité englobe des éléments qui ne lui appartiennent que partiellement. Par exemple, la rue *Au Maire* a été allongée en 1833 par une percée à l'Est. Pour autant, cette percée n'est pas un événement instantané mais un certain temps s'est déroulé entre la décision d'ouverture et la fin des travaux. La date donnée est donc une approximation de la véritable tranche temporelle qui correspond à la durée de l'opération.

Ambiguïté

L'ambiguïté est induite quand une entité appartient, intégralement ou en partie, à plusieurs catégories possibles. Pour De Runz (2008), un tel problème résulte soit d'un manque de spécificité de la définition de l'entité, soit d'un désaccord entre plusieurs définitions. Dans les deux cas, le problème peut être dû soit à des divergences entre des avis d'experts, ou bien directement par des incohérences dans les sources de données.

- Une entité présente un manque de spécificité lorsque plusieurs classes peuvent lui correspondre. La rue *Au Maire* est en partie construite sur un ancien cul-de-sac, parfois dénommé *Cul-de sac de Rome* ou *Passage de Rome*, bien que sa structure n'ait pas changé. Il est clair ici que deux classes de voies (cul-de-sac et passage) dont le sens est très différent peuvent s'appliquer à la même entité. Il faut remarquer que ce manque de spécificité, du moins lorsque l'on travaille avec des plans anciens, résulte en partie des différences d'échelles et

de niveaux d'interprétation des informations qu'ils contiennent.

- Enfin, une information sur une entité peut rentrer en conflit avec une autre. Dès lors, une même entité peut posséder plusieurs valeurs pour chacun de ses éléments constitutifs qui sont en désaccord. Plusieurs définitions d'une entité peuvent également être en conflit, par exemple lorsque les avis de deux experts divergent.

1.4.2.2 Incomplétudes

Enfin, De Runz (2008) spécifie deux autres types d'imperfections qui se différencient substantiellement des autres, à savoir les lacunes et les absences. Ces deux types sont des sous-catégories de la notion d'incomplétude abordée par Bouchon-Meunier (1995). L'absence désigne le fait qu'il manque des informations pour décrire complètement une entité. Par exemple, même si l'on sait que la rue *Au Maire* existait en 1280, on ne dispose pas de sources sur son emprise spatiale. Des trois éléments espace-temps-sémantique qui participent à la définition de l'entité rue, l'espace est dans ce cas absent. Les lacunes ajoutent une dimension hiérarchique à la notion d'incomplétude. En effet, une entité est lacunaire lorsqu'elle est composée de plusieurs entités dont certaines sont inconnues.

L'incomplétude est sans doute le type d'imperfection le plus répandu dans les informations géohistoriques. En effet, ces informations sont issues de sources partielles, éparses et régulièrement incomplètes. Ainsi, aucun plan ne décrit exhaustivement un espace, et chacun décrit de façon différente un même espace. Enfin, comme toute source d'information, les plans ne donnent qu'un aperçu d'un espace pendant une courte période. Entre ces périodes de temps, l'état de l'espace doit être supposé inconnu.

1.5 État de l'art des SIG utilisés en histoire

Diffusion de données historiques

Nous avons identifié la diffusion de données, c'est à dire la mise à disposition de données de manière large et systématique, comme un des facteurs importants de l'engouement des sciences sociales envers les SIG. Les données géographiques ainsi diffusées sont généralement associées à des plateformes de webmapping permettant d'accéder de façon plus rapide et plus intuitives aux données voulues. Étant donné le coût de développement d'un site web dédié à la diffusion de telles données, on y retrouve essentiellement des plateformes mises en place au niveau d'un État. C'est par exemple le cas des SIG historiques de Grande Bretagne (GBHGIS) (Southall et Aucott, 2009) et des États-Unis (NHGIS) (McMaster *et al.*, 2003) portant respectivement sur l'ensemble des deux pays. De tels SIG sont alors dédiés à la diffusion à grande échelle de données démographiques et sociales socio-économiques spatialisées, notamment les recensements du pays à différentes dates : ce sont des index géohistoriques ou *historical gazetteers*.

D'autres SIG visent d'avantage la diffusion de cartes anciennes et se constituent en atlas historiques en ligne. Ces outils mettent à disposition du public des collections de cartes historiques, généralement topographique et sous forme raster (Rumsey *et al.*, 2002). Le SIG historique ALPAGE (Noizet *et al.*, 2008) est un cas particulier de SIG dédié à la diffusion de données. En effet, les données diffusées par le SIG ALPAGE sont, pour la majorité, des informations spatiales vectorisées à partir de divers documents anciens. S'il ne s'agit pas d'un index géohistorique comparable aux SIG précédents, l'objectif d'ALPAGE ne se limite pas à la diffusion de cartes anciennes. La mise à disposition de données topographiques (cadastre napoléonien vectorisé, limites anciennes, etc.) permet de fournir, en plus d'informations sur un espace (en l'occurrence, Paris), des données manipulables par des outils SIG extérieurs dans des buts d'analyse spatiale ou de géocodage.

Représentation visuelle

La représentation visuelle de données historiques constitue également un des besoins de la recherche en histoire. Il n'est donc pas étonnant de retrouver des outils de visualisation cartographique dans la majorité des SIG historiques présentés dans le tableau 1.20. Si des outils de représentation visuelle sont systématiquement disponibles, la nature de ces outils dépend des objectifs du SIG : soit il s'agit de communiquer les résultats d'une étude historique, soit il s'agit de visualiser des données historiques dans un but de diffusion.

Différents moyens sont alors utilisés. Lorsqu'il s'agit de communiquer des résultats de recherche, le SIG historique peut simplement mettre à disposition des cartes créées pour les besoins de l'étude, comme par exemple pour le projet *Valley of the Shadow* (Ayers, 2003). Plus souvent, la mise en place d'un outil de *webmapping* permet de communiquer au sein d'un seul outil l'ensemble des données créées ou assemblées au cours de l'étude. La diffusion de données s'appuie également largement sur de tels outils, permettant un accès plus intuitif à des données spatiales.

Bien qu'il s'agisse également d'un index géohistorique, l'outil en ligne GAPVIS (Isaksen *et al.*, 2012) vise un but différent. L'objectif de ce SIG historique est de cartographier les lieux et les relations entre lieux décrits au sein de récits anciens. À partir de textes de livres anciens issus de *Google Books*, les lieux sont automatiquement détectés et géocodés à l'aide de l'index géographiques web externes Pléiades²².

Gestion des imperfections

Nous l'avons vu, les imperfections des informations géohistoriques sont nombreuses et se trouvent à toutes les étapes de l'intégration de données dans un SIG. la gestion de ces imperfections devient primordiale lorsque l'on traite de données fortement imparfaites dont les défauts peuvent mener à une lecture erronée. Plus radicalement, de trop fortes imperfections peuvent s'accumuler pour rendre le phénomène retranscrit illisible. Pourtant, rares sont les SIG historiques qui modélisent explicitement les différentes imperfections des données.

22. <http://pleiades.stoa.org/>

La majeure partie du temps, les données font l'objet de pré-traitements de façon à réduire ou éliminer ces imperfections. Les données raster -généralement des cartes anciennes- sont géoréférencées en amont de leur mise à disposition. Cette opération aborde par nature certaines imperfections des plans, notamment les erreurs géométriques qui empêchent un recalage parfait de ces plans sur un référentiel récent. Les lacunes et les choix de représentations de ces sources (échelle homogène ou non, éléments non représentés, vues à projection multiples, etc.) peuvent également affecter le géoréférencement de ces sources. Pourtant, rares sont les SIG qui prennent en compte de telles imperfections. Lorsqu'il s'agit de cartes, cette prise en compte passe généralement par la communication des déformations et erreurs résultantes ainsi que les choix des méthodes de géoréférencement adoptées dans des métadonnées. La plateforme de diffusion de cartes anciennes *David Rumsey Map Collection* (Rumsey *et al.*, 2002) par exemple met à disposition du public de très nombreuses cartes anciennes géoréférencées, sur le monde entier. Pourtant, les informations portant sur le géoréférencement des cartes diffusées ne sont pas disponibles, malgré la présence de nombreuses métadonnées sur chaque carte.

À l'opposé, le SIG *Historical Counties* (Kelly, 2002) fournit un certain nombre d'informations sur les imperfections des données qu'il manipule au travers de métadonnées. L'objectif de ce SIG est de mettre à disposition du public l'ensemble des données vecteur décrivant les évolutions des limites des comtés américains, depuis le début du XVII^e siècle jusqu'à l'année 2000. L'ensemble des frontières des comtés ont été saisies à partir d'un référentiel récent qui varie selon le comté, chaque frontière étant par la suite située temporellement dans une base de données par ses dates d'apparition et de disparition. À chaque comté ainsi vectorisé est associé un ensemble de métadonnées basées sur le schéma de métadonnées Dublin Core (Dublin Core Metadata Initiative, 2012), décrivant notamment la précision de localisation des données par rapport aux données de références utilisées lors de la saisie, mais aussi une description fine des différentes opérations et des différentes sources historiques ayant mené à la constitution complète des frontières du comté.

(Noizet *et al.*, 2008) constitue ici également un cas à part. Cette plateforme, diffusant entre autre des données vectorisées sur des plans de villes anciens, fournit diverses métadonnées décrivant les imperfections des plans et de leur contenu. Ainsi, pour chaque couche vectorisée, les erreurs résiduelles du géoréférencement sont disponibles dans les métadonnées associées à la couche. De plus, une attention particulière est portée au contenu des plans vectorisés. Ainsi, aux îlots du cadastre napoléonien de Paris sont associés un indice de fiabilité décrivant la qualité de la représentation de l'îlot dans le plan, et indiquant les éventuels problèmes de géoréférencement.

Si les imperfections des données peuvent être représentées au sein d'un méta-modèle, il est également envisageable de fournir des outils de visualisation de ces imperfections. Cependant, peu de SIG historiques proposent une telle visualisation. Nous retenons ici l'exemple du projet *Mapping the jewish communities of the byzantine empire* (de Lange *et al.*, 2013) dont l'objectif est de cartographier les communau-

tés juives dans la zone géographique de l'empire byzantin. Les sources historiques, souvent textuelles, qui ont permis l'identification de ces communautés sont un cas de sources spatiales implicites. Dès lors, les auteurs ont dû reconstituer les emplacements de chacune des communautés, impliquant donc un haut niveau d'incertitude quant au placement de chaque peuplement. Afin de traduire ces incertitudes, le choix a été fait de mettre en place un code graphique décrivant le niveau d'incertitudes sur chaque élément de l'information géohistorique : la localisation de chaque communauté, sa période d'existence et sa dénomination (voir figure 1.17). Ce SIG fournit en outre des informations sur les lacunes des sources dans les métadonnées associées à chaque communauté identifiées.

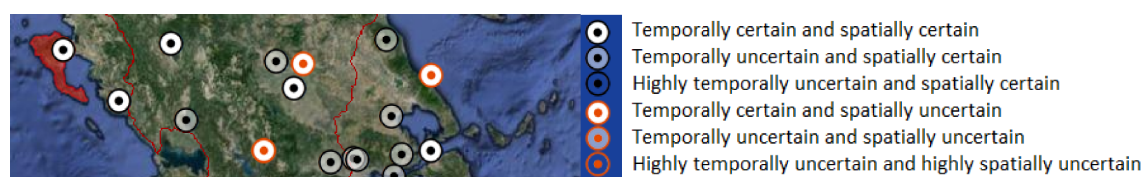


FIGURE 1.17 – Extrait du SIG *Mapping Jewish Communities* (de Lange *et al.*, 2013) et légende associée. Les lieux de peuplement des communautés juives de l'empire Bizantin sont symbolisés en fonction des incertitudes associées à leurs composantes sémantiques, spatiales et temporelles.

Gestion des dynamiques spatiales

La représentation, visuelle ou bien dans le modèle de données, des dynamiques de l'espace constitue un besoin crucial des historiens travaillant sur des problèmes spatialisés. Il paraît donc naturel que les SIG historiques fournissent des outils permettant une telle représentation. Cependant, force est de constater qu'aucun systématisme n'existe sur les méthodes de représentation ou le niveau de considérations pour ces dynamiques.

La plupart du temps, les SIG historiques se construisent autour de modèles spatiaux et temporels (voir le paragraphe 1.1.2) dans lequel chaque donnée -carte, couche vecteur, etc.- est associée à une date. Cette méthode d'association est nommée *date stamping* (Gregory et Healey, 2007), et, si elle constitue un moyen simple et efficace pour situer temporellement des données historiques, elle ne permet pas de modéliser une connaissance sur leurs évolutions. Souvent, cette représentation est associée à des curseurs permettant de naviguer sur la ligne du temps, les données correspondant à la date sélectionnée s'affichant alors. La représentation des dynamiques de l'espace est donc implicite et doit être recomposée mentalement par l'utilisateur en *déroulant le temps*. C'est cette représentation qui se trouve au cœur de nombreux SIG historiques (Gordon, 2008; Kelly, 2002; de Lange *et al.*, 2013; Shnayder, 2010; Isaksen *et al.*, 2012; Desender, 2009) ou encore l'outil TimeMap (Johnson, 2004).

Quelques SIG historiques ont développé un modèle de données visant à représenter les dynamiques de données spatiales. Le SIG *China GIS* (Bol, 2007) en

particulier développe un modèle de données se fondant sur la représentation des filiations entre zones administratives existantes à différentes époques. Représentées par des surfaces vectorisées, chaque unité administrative peut être associée à des successeurs dans le temps par des liens de filiations décrivant le type de transformations. Ainsi, une unité administrative existant à un temps $t = 1$ donné correspondant à l'emprise spatiale de trois unités administratives à $t = 2$ créera une relation de scission, la première unité ayant été séparé en deux plus petites (voir la figure 1.18).

Enfin, certains SIG historiques, également axés sur les transformations de zones administratives, modélisent les transformations de l'espace en décrivant les modifications des frontières. Les frontières, généralement considérées à partir d'une période initiale, sont dans un premier temps représentées et associées à une tranche temporelle. Lorsqu'une modification affecte une frontière, celle ci est découpée en plusieurs parties. Les parts ne changeant pas voient leur période temporelle agrandie, tandis que les parties modifiées donnent naissance à de nouvelles frontières. Finalement, les frontières deviennent un amalgame d'*atomes* spatio-temporels, qui sont associés à chaque date pour former des frontières complètes. Les SIG *Historical Counties* (Kelly, 2002) et *China GIS* (Bol, 2007) mettent notamment en place de telles architectures.

Notons dès à présent que ces modèles constituent des instanciations, dans un cadre historique, de modèles de données spatio-temporels existants dans la littérature. Ces différents modèles seront abordés en détails dans le chapitre 3.

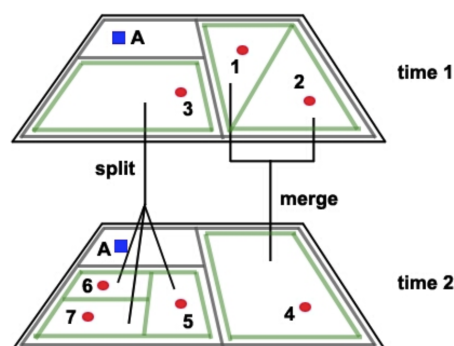


FIGURE 1.18 – Liens de filiations entre unités administratives dans le SIG *China GIS* Bol (2007). Issue de Berman (2013).

Intégration de données historiques

Au delà de la diffusion de données historiques, les historiens voient dans les SIG un moyen de partage d'informations. Afin de mettre en place des systèmes ouverts et évolutifs, il est donc important de prévoir des méthodes permettant d'ajouter de nouvelles données. Puisque le passage entre les sources de données historiques et les bases de données géographiques nécessitent de nombreuses opérations, la création de méthodes permettant d'intégrer de nouvelles informations peut devenir complexe. Nous entendons ici le terme *intégrations de données* au sens de Li et Goodchild

(2011a), c'est à dire la combinaison de deux ou plus sources d'informations permettant d'obtenir une informations impossible à acquérir avec une source seule. Ici, l'intégration de données concerne deux éléments :

1. la mise en relation de données spatiales implicites de façon à leur attribuer une localisation dans l'espace ;
2. la combinaison de données spatiales et d'informations sur ses transformations de façon à créer des données spatio-temporelles.

Le premier aspect de l'intégration s'illustre notamment au travers de GAP-VIS (Isaksen *et al.*, 2012), dont l'objectif est de cartographier les récits contenus dans des livres anciens (antiquité romaine en particulier). Ainsi, à partir de textes anciens, les noms de lieux et les relations entre lieux sont extraits de façon automatique, puis associés à une localisation dans l'espace grâce à l'index géographiques en ligne Pléiades. De même, la plateforme *Mapping London's Past* (Shoemaker *et al.*, 2011) possède un outil permettant de géocoder automatiquement des noms de lieux dans Londres ainsi que des outils de post-traitement manuel et de vérification des données géocodées. Notons également que le SIG historique de Grande Bretagne subit actuellement un profond remaniement de son modèle, s'intéressant notamment à l'intégration d'informations qualitatives portant sur des lieux décrits au sein de textes (Southall, 2012).

Le *China GIS* illustre pour sa part le second point de l'intégration de données géohistoriques. Les données manipulées par le SIG sont un ensemble de zones administratives de la Chine ancienne, représentées sous la forme de polygones à différentes périodes historiques. À chacune des unités administratives est associé une sémantique (son nom et ses liens hiérarchiques) ainsi que la période temporelle de son existence. Les différentes données sont alors tout à fait asynchrones, chacune étant associée à une période propre. Afin de permettre la représentation de l'ensemble des unités administratives à une date choisie tout en évitant la réplication des données, le choix à été fait de fusionner les données au sein d'une couche unique lors d'une requête temporelle. Ainsi, lorsqu'une représentation à un moment donné est demandée, le SIG agrège les différentes unités administratives dont la période d'existence comprend la date demandée, tel qu'illustré dans la figure 1.19. Les transformations des unités administratives peuvent alors être obtenues en fusionnant plusieurs couches à des dates différents. Leur recouvre-

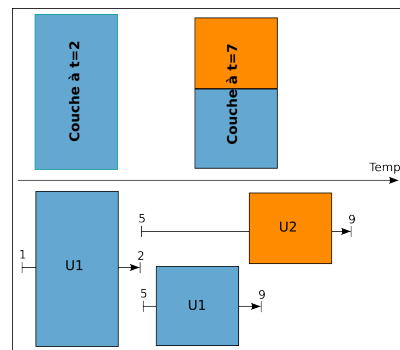


FIGURE 1.19 – Modèle d'objets géohistoriques asynchrones et couches fusionnées dans le *China GIS*

à une période propre. Afin de permettre la représentation de l'ensemble des unités administratives à une date choisie tout en évitant la réplication des données, le choix à été fait de fusionner les données au sein d'une couche unique lors d'une requête temporelle. Ainsi, lorsqu'une représentation à un moment donné est demandée, le SIG agrège les différentes unités administratives dont la période d'existence comprend la date demandée, tel qu'illustré dans la figure 1.19. Les transformations des unités administratives peuvent alors être obtenues en fusionnant plusieurs couches à des dates différents. Leur recouvre-

ment spatial permet d'identifier les liens de filiations existants entre les différentes données.

Enfin, puisqu'il s'agit pour la majorité de plateformes web, des méthodes d'intégration de données pourraient permettre d'ouvrir le SIG vers l'extérieur en le transformant en SIG collaboratif. Pourtant, nous avons constaté qu'aucun n'a fait ce choix. L'outil TimeMap (Johnson, 2004) permet à l'utilisateur d'ajouter ses propres informations, mais il s'agit d'avantage d'un outil de création de cartes dynamiques que d'un système d'information géographique. Finalement, le SIG en ligne *HistoAtlas* (Desender, 2009) est pour sa part entièrement basé sur un processus collaboratif, la plus grande part de son contenu étant créé par les utilisateurs. Ainsi, toute personne peut ajouter des objets géographiques (pays ville, événements, etc.), les situer dans le temps, créer des hiérarchies entre objets et des cartes. Finalement, les données et les cartes ainsi créées sont accessibles et modifiables par d'autres utilisateurs.

Bilan sur les SIG historiques existants

Nous avons évoqué jusqu'ici les différentes stratégies adoptées par les SIG historiques pour répondre aux demandes des historiens. Dans cette étude rapide, on peut constater que les SIG existants répondent à la majorité de ces besoins, mais souvent de manière incomplète. Ainsi, la représentation visuelle des incertitudes ou des dynamiques l'emporte souvent sur leur prise en compte dans le modèle de données associé au SIG. L'objectif premier de ces outils étant la diffusion de données ou de résultats d'études, il n'est pas étonnant que l'accent soit mis en priorité sur la représentation visuelle des données. De plus, les outils de webmapping existants²³ constituent des outils puissants et aboutis qui permettent aux historiens, souvent utilisateurs finaux des outils SIG, de mettre en place des plateformes web avec une relative facilité. Cependant, de tels outils sont inadaptés pour la représentation de données imparfaites ou la modélisation des dynamiques de l'espace. En conséquence, ils ne fournissent aucun moyen de modéliser, représenter ou intégrer des données spatiales comportant ces deux aspects. On peut alors penser que cette inaptitude freine largement la mise en place de SIG historiques répondant de manière complète aux besoins historiques.

Afin d'explicitier cette constatation, nous résumons les différents points abordés dans le tableau 1.20. Pour chaque système abordé dans les paragraphes précédents, nous avons représenté sa capacité de réponse aux demandes des historiens. Chaque demande est décomposée en quelques stratégies adoptées *de facto* par les SIG existants. Un violet clair signifie que le SIG ne répond pas au point correspondant, tandis que le violet le plus foncé signifie que le SIG possède une fonctionnalité répondant totalement au critère. Enfin, une couleur intermédiaire est présente dans le cas où le SIG répond de façon partielle à un critère, la grande diversité de ces outils empêchant parfois un classement clair dans une catégorie précise. Par exemple, si le *China GIS* aborde la question de l'intégration de données par la méthode décrite précédemment, il ne propose pas de véritable processus d'intégration permettant d'ajouter de nouvelles informations de façon automatique et donc extensibles à d'autres données.

23. Tels Geoserver (<http://geoserver.org>) et Mapserver(<http://mapserver.org>)

1.6 Positionnement

Constatations

Au travers des sections précédentes, nous avons vu que la reconstitution précise de l'espace ancien constitue une ressource de première importance dans le cadre d'un travail historique sur la ville. Elle agit comme un pivot pour l'étude de données sociales car il permet de situer les individus, les activités et les événements au sein d'un référentiel commun. Mais il peut être également lui-même objet d'étude lorsqu'il est considéré dans une relation systémique avec les activités humaines, lorsqu'il est pris comme composante d'un territoire. Dans les deux cas, s'appuyer sur l'espace pour étudier un phénomène social par nature dynamique impose de considérer un espace également dynamique, dont les transformations dans le temps se lient avec les activités humaines.

Les systèmes d'informations géographiques, par leur capacité à manipuler, stocker et représenter une information spatiale ont attiré les sciences sociales, en particulier les historiens, au cours des dix dernières années. Pour un travail historique, les SIG sont dès lors de puissants outils permettant de centraliser, de partager, d'analyser et de cartographier des phénomènes sociaux et spatiaux s'étant déroulés dans le passé. Ils permettent de fournir des représentations visuelles, mais aussi de constituer des bases de données spatiales sur les formes des espaces passés de grande taille.

Lorsqu'il s'agit de la ville, la constitution de représentations et de bases de données globales dépendent de nombreuses sources historiques extrêmement hétérogènes. Les cartes anciennes, en particulier les plans de villes forment le socle fondamental de la reconstitution d'un espace ancien. Très diverses, ces sources, tout comme l'information spatiale qu'elles portent sont imparfaites. Ainsi, la fiabilité de chaque source historique doit être questionnée avant d'être utilisée dans un travail historique, les informations qu'elle portent pouvant être incertaines, ambiguës ou lacunaires.

Si la prise en compte de ces imperfections fait partie intégrante de la pratique quotidienne d'un chercheur, habitué à confronter chaque information à d'autres sources en les croisant et en analysant leur consistance, la capacité des SIG à manipuler une telle information doit être questionnée. Au travers de l'exploration de différents SIG dédiés à l'étude historique, nous avons clairement fait apparaître la faible capacité de ces systèmes à manipuler, modéliser et représenter des données spatiales imparfaites. Nous l'avons vu, la prise en compte des imperfections de l'information se limite la plupart du temps à la minimisation des déformations des cartes géoréférencées. Seul les SIG *ALPAGE* et *Historical Counties* tentent de modéliser l'incertitude de leurs données au sein d'un schéma de métadonnées. Quant à la représentation visuelle des incertitudes, seul le SIG *Mapping Jewish Communities* (de Lange *et al.*, 2013) propose une symbolisation spécifique de l'incertitude des données qu'il cartographie.

Enfin, la modélisation des dynamiques de l'espace manque également dans la majorité des SIG existants, malgré l'importance de ce besoin parmi la communauté des historiens. En effet, bien que cette dimension apparaisse comme essentielle pour analyser les transformations connues au cours du temps par un lieu et une société, seuls les auteurs du *China GIS* (Bol, 2007) modélisent explicitement les transformations des données spatiales qu'ils manipulent. Dans le cas du SIG *Historical Counties* et *TimeMap*, les transformations des objets spatiaux ne sont pas modélisées explicitement. En effet, seules sont stockées des dates de début et de fin d'existence de chaque objet, la représentation des dynamiques étant dès lors uniquement visuelle grâce à la production de cartographies dynamiques. Notons cependant que pour *Historical Counties*, les changements sont tout de même décrits dans les métadonnées.

Finalement, on constate au travers du tableau 1.20 que pour la majorité des SIG, l'intégration des données est effectuée soit *a priori*, soit de façon manuelle. Plus encore, les seuls systèmes qui possèdent des outils d'intégration automatique des données historiques sont dédiés à la localisation de lieux ou d'individus, c'est à dire à au géocodage de l'information. Ainsi, GAPVIS s'appuie sur des index géographiques en ligne pour localiser des lieux anciens, tandis que le SIG *Locating London Past* utilise des banques de données d'adresses pour placer des individus dans l'espace.

L'ordonnancement chronologique proposé dans le tableau 1.20 ne fait pas apparaître de tendances particulières dans l'évolution de ces systèmes, sinon par le développement des outils de visualisation de données en ligne. Si une considération pour les imperfections des données semble apparaître à partir de la publication du *China GIS*, la modélisation des dynamiques de transformation des objets étudiés est rarement traitée, la cartographie dynamique restant l'outil d'exploration spatio-temporel privilégié. Il faut cependant rester prudent : le tri temporel est effectué sur les dernières versions de chacun des SIG, certaines versions anciennes peuvent présenter des fonctionnalités complexes. C'est par exemple le cas pour *Historical Countier*, dont la création des bases de données spatiales a commencé en 1987²⁴.

Hypothèses et objectif

Nous nous plaçons dans le cadre de la modélisation de l'espace urbain ancien au sein d'un SIG historique. Nous avons identifié dans les sections 1.3 et 1.3 deux demandes importantes des historiens adressées à de tels outils :

- La possibilité de centraliser, croiser et diffuser des données et des sources géohistoriques.
- La modélisation des dynamiques de l'espace en vue de leur analyse.

Au travers de l'exploration des SIG historiques existants, nous avons constaté qu'ils répondent à l'heure actuelle de façon satisfaisante à la première demande, notamment grâce aux initiatives de création de SIG en ligne de niveau national tels *China GIS*, *NHGIS* ou *GBHGIS*.

Le deuxième point cependant fait cruellement défaut au sein des outils et systèmes existants à l'heure actuelle. S'il est communément admis (Gregory et Healey,

24. <http://publications.newberry.org/ahcbp/project.html>

2007; Knowles et Hillier, 2008) que ce manque est principalement dû à l'incapacité des outils grand public à manipuler des données imparfaites et spatio-temporelles, force est de constater que le développement des systèmes plus spécialisés n'a pas abouti à la proposition d'outils ou de modèles adaptés aux besoins de l'histoire. En effet, les SIG historiques restent aujourd'hui dédiés à des applications spécifiques et ne sont pas destinés à être utilisés comme outils dans d'autres études historiques.

On peut alors penser que l'absence d'un modèle répondant au second besoin des historiens provient d'un manque de méthodes dédiées à la constitution de bases de données spatio-temporelles à partir de données géohistoriques issues de sources hétérogènes. En effet, nous avons vu que dans la majorité des cas, les données spatio-temporelles sont créées de façon manuelle à partir de diverses sources, qu'elles aient été saisies en amont de la construction du SIG ou ajoutées au fur et à mesure. Tout comme l'indique Arnaud (2008), la reconstitution des dynamiques spatiales à partir uniquement de sources historiques agissant comme des *tranches temporelles* est une tâche fastidieuse dont la complexité est extrême lorsque les sources sont imparfaites.

La mise en place d'un processus d'intégration de données géohistoriques visant à constituer des données spatio-temporelles peut permettre d'augmenter la généralité d'un SIG historique. L'automatisation de ce processus est donc un atout important. Cependant, nous pensons que la complexité des sources géohistoriques ainsi que les connaissances mises à l'œuvre pour comprendre, analyser et critiquer une source rendent peu pertinente la mise en place de processus d'intégration entièrement automatisés.

Comme nous l'avons vu dans la section 1.1, la reconstruction d'un espace ancien permet de produire des données géohistoriques à la fois spatiales et temporelles. Afin d'obtenir une représentation des dynamiques de cet espace, il est nécessaire d'opérer la mise en relation des différents objets porteurs d'informations spatiales pour des temporalités différentes pour constituer des bases de données géohistoriques de nature proprement spatio-temporelles. Une fois ces données constituées, leurs transformations peuvent être analysées.

Nous avons également souligné en section 1.2.3 l'importance des plans lorsqu'il s'agit d'étudier un espace ancien. Dans le cas des villes plus particulièrement, et pour des périodes récentes, ces sources fournissent la partie la plus significative des informations sur leurs structures et leur organisation. En nous fondant sur ce constat, nous avons donc choisi de concentrer notre travail sur le traitement des informations spatiales issues de plans de villes anciens. Plus précisément, dans le cadre de cette thèse, l'espace considéré sera la ville de Paris entre les XVIII^e et XX^e siècles, en raison de la richesse des ressources cartographiques décrivant la ville, mais également de l'intensité des transformations qu'elle connaît en l'espace de deux siècles (voir le chapitre 2).

En résumant, dans notre travail nous nous sommes fixé trois objectifs principaux :

- Constituer une base de données spatiale et temporelle de données vectorielles extraites de plans de villes anciens.
- Proposer aux historiens une méthode semi-automatique pour la constitution de bases de données spatio-temporelle, à partir de l'analyse, du traitement et de la mise en relation de sources historiques à contenu spatial.
- Modéliser et de prendre en compte, au sein des différentes étapes de l'intégration des données géohistoriques, les différentes imperfections qui entachent l'information manipulée.

1.7 Conclusion

Dans ce chapitre, nous avons tout d'abord identifié les principales fonctionnalités qu'un SIG historique doit posséder pour être adapté aux besoins des chercheurs en histoire. Pour cela, nous avons exploré comment l'espace géographique étant actuellement utilisé dans différents travaux de recherche en histoire urbaine. Nous avons vu la nécessité de visualiser, analyser et manipuler les dynamiques de transformation des objets géographiques urbains. En effet, les phénomènes sociaux étudiés évoluent conjointement à l'espace dans lequel ils se déroulent. Ils interagissent également avec lui, rendant ainsi leur analyse complète seulement possible si l'espace est considéré dynamique. Sur ce point, nous avons listé les principaux SIG historiques existants et constaté qu'ils ne répondent pas à ce besoin. Ceux-ci sont en réalité statiques, peu extensibles et servent principalement de plateforme de diffusion de données brutes et de résultats de travaux. Ils ne fournissent en général pas d'outils d'analyse de ces données et ne permettent pas de manipuler et d'analyser des objets géographiques en transformation dans le temps. Ces manques sont notamment dus aux complexités des sources historiques, imparfaites et difficiles à intégrer car demandant d'importants traitements (numérisation, vectorisation, etc.).

De cette étude, nous avons dégagé l'objectif principal de cette thèse : mettre en place une approche permettant de constituer des bases de données spatio-temporelles décrivant les transformations de l'espace urbain à partir de sources cartographiques anciennes.

Notre cas d'application est la ville de Paris aux XVIII^e et XIX^e siècles. Nous avons pointé dans ce chapitre la nécessité d'une connaissance fine des sources utilisées si l'on veut utiliser sans risque d'erreurs les informations qu'elles contiennent. Nous procédons à cette analyse dans le prochain chapitre sur cinq sources cartographiques parisiennes que nous utiliserons pour expérimenter et valider l'approche proposée au long de ce travail.

	Diffusion de données		Visualisation		Incertitudes		Dynamiques		Intégration		Date de publication
	Cartes Données Spatiales	Carto. en ligne	Résultats de recherche	Modèle Métamodèle	Rep. visuelle	Rep. visuelle Modèle	<i>A priori</i>	Méthode d'intégration	Collaboratif		
SIG Historique											
<i>TimeMap</i>	■	■	■	■	■	■	■	■	■	■	2004
<i>Mapping Decline</i>	■	■	■	■	■	■	■	■	■	■	2008
<i>HistoAtlas</i>	■	■	■	■	■	■	■	■	■	■	2009
<i>GBHGIS</i>	■	■	■	■	■	■	■	■	■	■	2009
<i>NHGIS</i>	■	■	■	■	■	■	■	■	■	■	2010
<i>CHGIS</i>	■	■	■	■	■	■	■	■	■	■	2010
<i>Alpage</i>	■	■	■	■	■	■	■	■	■	■	2010
<i>Locating London past</i>	■	■	■	■	■	■	■	■	■	■	2011
<i>Valley of the shadow</i>	■	■	■	■	■	■	■	■	■	■	2011
<i>MJCB</i>	■	■	■	■	■	■	■	■	■	■	2011
<i>GAPVIS</i>	■	■	■	■	■	■	■	■	■	■	2011
<i>Historical counties</i>	■	■	■	■	■	■	■	■	■	■	2012
<i>Spatial History</i>	■	■	■	■	■	■	■	■	■	■	2013
<i>David Rumsey Collection</i>	■	■	■	■	■	■	■	■	■	■	2013
<i>Geoportail</i>	■	■	■	■	■	■	■	■	■	■	2013

FIGURE 1.20 – Tableau récapitulatif des fonctionnalités de 15 SIG historiques existants. ■ : Non prise en compte. ■ : Prise en compte partielle ou traitement manuel. ■ : Prise en compte totale.

Chapitre 2

Sources cartographiques sur Paris

Objectifs :

- Présenter les sources topographiques parisiennes utilisées.
- Identifier les hétérogénéités des sources et les liens d'héritages existants entre elles.
- Déterminer de façon qualitative les temporalités valides des sources.
- Explorer la mutation de d'administration de la Seine à travers les plans de Paris.

Sommaire

2.1	Atlas de Verniquet	67
2.1.1	Levé et structure	67
2.1.2	Contenu du plan	69
2.2	Plan de Maire	69
2.2.1	Structure du plan	70
2.2.2	Contenu du plan	71
2.3	Cadaastre de Vasserot	73
2.3.1	Levé et structure du plan	74
2.3.2	Contenu du plan	74
2.4	Atlas de Jacoubet	75
2.4.1	Levé et structure du plan	76
2.4.2	Contenu du plan	76
2.5	Atlas municipal	77
2.5.1	Structure du plan	78
2.5.2	Contenu	79
2.6	L'atlas de Jacoubet : des architectes voyers à l'administration de la Seine	80
2.6.1	De la servitude d'alignement à l'Atlas Général	81
2.6.2	Création et première datation	84
2.6.3	Datation fine	88
2.6.4	Jacoubet et l'administration parisienne : l'insertion de l'atlas gé- néral dans la gestion de la Ville	94
2.6.5	Notes et remarques	102
2.7	Conclusion	107

Traiter des plans anciens au sein d'un SIG implique nécessairement qu'ils soient proches des plans topographiques contemporains pour lesquels ont été développés les outils et les méthodes de traitement actuels. Cette proximité est tout d'abord cartographique : la vue doit être verticale, les thèmes cartographiques semblables. Elle doit être aussi géométrique, dans la précision du positionnement des objets dans le plan de la carte : il faut donc des plans marqués par une forte précision, présentant de faibles déformations géométriques. Les plans en perspective, majoritaires avant le XVIII^e siècle, sont donc difficilement utilisables par un SIG à deux dimensions¹ Balletti (2000), ayant référencé un plan de la ville de Venise en plaçant un ensemble de points d'amer sur des bâtiments dessinés en perspective, a conclu à une difficulté de comparaison du résultat avec des données planimétriques, et ce, malgré un géoréférencement satisfaisant. Plus généralement, de telles vues sont davantage utilisées dans le cadre de SIG à trois dimensions ou pour la reconstitution d'édifices anciens en archéologie (Stefani, 2010). C'est par exemple le cas du *Virtual Lodeum* (Pfeiffer *et al.*, 2013), SIG 3D de la ville de Liège constitué à partir d'une maquette de 1730, dans lequel des gravures anciennes de bâtiments peuvent être projetées sur les modèles 3D correspondants.

Les transformations les plus importantes de l'espace parisien que l'on peut aisément documenter se situent durant les XVIII^e, XIX^e et XX^e siècles.

La population parisienne passe ainsi d'environ 640.000 âmes en 1793 (Landry, 1935) à 600.000 en 1811, dépasse le million en 1846 pour atteindre 1.200.000 en 1856 (Chevalier, 1952) dans les mêmes limites. En 1891 la ville compte plus de 2.440.000 habitants, soit près de quatre fois la population cent ans plus tôt. La ville se transforme en profondeur et à plusieurs reprises : la Révolution amène la vente des anciennes propriétés ecclésiastiques et libère de larges espaces lotissables, tandis que le Consulat et le Premier Empire mettent en place des politiques de grand travaux : canal de l'Ourcq et réservoirs, abattoirs, quais, etc. Aux grandes opérations de lotissement de Grenelle et Plaisance au sud, Passy à l'ouest ou encore Poissonnière et Batignolles au nord succèdent les premières grandes percées : la poursuite de la rue de Rivoli au travers du centre parisien ou la rue Rambuteau sont ouvertes. Au milieu du XIX^e siècle c'est le centre parisien qui subit les politiques hygiénistes des autorités centrales ; les rues sont élargies, rendues rectilignes, les impasses deviennent des rues. Survient ensuite le rattachement des communes limitrophes qui agrandit Paris de 4.365 hectares, puis les opérations haussmanniennes qui poursuivent la transformation de la ville. Après 1885, Paris ne connaîtra plus de transformations majeures, sinon dans le quartier de la Villette en raison du développement des industries (Bourillon, 2000).

Dans ce cadre, nous avons choisi de nous concentrer sur la période allant de la Révolution Française à la fin du XIX^e siècle. Ainsi, nous couvrons une période marquée par des modifications majeures mais aussi par une riche production de plans topographiques. En effet, la rapidité des transformations urbaine s'accompagne d'une profusion de plans souvent réalisés par des architectes.

1. On parle de SIG 2D pour désigner les outils classiques destinés à la cartographie, en opposition aux SIG 3D qui permettent de représenter des données géographiques en volume.

Simple guides de la ville, la plupart de ces plans sont à une échelle trop grande et sont trop peu précis pour être utilisables comme sources principales d'information sur l'espace parisien. Certains d'entre eux se distinguent cependant par une très grande précision et une échelle fine, entre 1/2.000^e et 1/10.000^e. Ils sont pour la plupart liés aux instances du pouvoir parisien et ils étaient destinés à être de véritables outils d'aménagement de la ville, ce qui permet de les considérer comme les sources les plus précises et fiables sur la topographie de Paris. Ainsi, nous avons sélectionnés cinq plans qui nous permettent de couvrir la totalité de la période considérée :

- l'*Atlas National de la ville de Paris* par Edmé Verniquet, publié entre 1791 et 1799. On le désignera par souci de simplicité *atlas de Verniquet*,
- la *Topographie de Paris*, plan réalisé par Nicolas Maire et publié en 1808. Il sera nommé *plan de Maire*,
- le *cadastre de Paris par îlots* de Philibert Vasserot, commencé en 1810 et terminé en 1836 (1855 pour certains îlots). On l'appellera *Cadastre de Vasserot*,
- l'*Atlas Général de la ville et des faubourgs de Paris* de Théodore Jacoubet, publié en 1836. Il sera nommé *atlas de Jacoubet*,
- l'*Atlas municipal des vingt arrondissements de la ville de Paris*, plan officiel de la préfecture de la Seine, édition de 1888. Il sera désigné comme un *atlas municipal*.

Parmi ces plans, seul le dernier est officiellement issu des services municipaux, tous les autres étant le produit d'entreprises individuelles². Nous verrons au cours de ce chapitre qu'il s'agit bien d'une caractéristique propre de cette période qui marque la nature de ces sources aussi bien du point de vue technique que de leur utilisation dans le cadre de la politique urbaine de Paris. Issus d'un travail individuel, ces plans étaient levés sur des périodes souvent longues. Ils étaient aussi objets d'un marché de vente, échange, reprise et diffusion qui impactait non seulement les formes des uns sur les autres mais induisant aussi la reproduction non contrôlée de leurs erreurs et imprécisions. Dans plusieurs cas, une analyse attentive, permet en effet de voir comment, dans une même planche, architectes et géomètres réutilisaient des levés ou des représentations cartographiques antérieures, conservant alors leurs erreurs et imprécisions. Si ces aspects sont connus des historiens, ils ont été rarement considérés dans la construction des SIG historiques existants qui acceptent de les intégrer sans un véritable travail critique. Au cours de ce chapitre, nous essayerons donc d'analyser qualitativement le contexte, la structure et le contenu de chaque plan, de façon à déterminer plus précisément leur nature, notamment en ce qui concerne les temps de validité, aspect qui est très souvent difficile de déterminer, comme nous avons vu au cours du chapitre précédent.

Une telle démarche s'est révélée particulièrement fructueuse. Ainsi nous avons pu démontrer que, pour l'ensemble de ces plans, les dates officielles de publication ne reflètent jamais la réalité de terrain des planches qui les composent. Emprunts, reprises, nouveaux levés et projets d'ouverture se superposent de manière souvent contradictoire en restituant surtout l'image d'un plan comme d'un millefeuille qui accumule les fragments locaux d'une ville qui change sans cesse. La complexité de ce tramage temporel est donc avant tout le résultat presque inévitable induit par les

2. Si le projet est à l'initiative d'une seule personne -architecte, commissaire voyer, etc.-, des équipes de taille variée travailleront à la réalisation de chaque plan. Plus de 40 géomètres participeront par exemple à la réalisation de l'atlas de Verniquet.

techniques de travail de l'époque. Mais nous avons pu voir qu'elle révèle aussi une dynamique plus complexe donnée par le changement des formes de production de l'objet cartographique qui passent de la dimension de l'entreprise privée à celle du public. Si Edmé Verniquet avait travaillé sur son plan en tant que géomètre privé commissionné par l'administration préfectorale, Théodore Jacoubet produira le sien en tant que fonctionnaire du Bureau des Plans de la ville de Paris. Dans les plans levés au cours du XIX^e siècle, nous trouvons donc aussi la trace de nombreuses tensions qui marquent ce processus. Tout au long du siècle les intérêts privés et intérêts publics se superposent de manière souvent contradictoire en impactant très souvent sur la qualité et même la véracité des levés. Au centre de cette période et de ces changements, l'analyse de la figure de Théodore Jacoubet et de son travail permet de mesurer très précisément la nature et l'importance de ces déterminations.

2.1 Atlas de Verniquet

L'*Atlas National de la Ville de Paris*³ est à la fois l'œuvre majeure de son auteur, l'architecte Edmé Verniquet, et la matrice de la plupart des plans parisiens du XIX^e siècle. Terminé en 1791, cet atlas composé de 72 feuilles à l'échelle de 1/1.800 environ (à l'origine, le plan est en toises à une échelle d'une demi-ligne par toise) est un portrait de la ville à l'aube de la Révolution. La précision extrême du plan et le style dépouillé qui lui sont caractéristiques feront de lui la base de la plupart des grands plans de Paris pendant de nombreuses années et un outil convoité par l'administration de la Seine. Ce n'est pourtant ni le plan le plus riche d'informations, ni le plus novateur de son époque. En 1730 déjà des cartographes tel Delagrive s'étaient attachés à concevoir des plans de Paris dits *géométraux* où la précision géométrique était l'objectif principal. On connaît aujourd'hui l'ampleur et les détails de la tâche de Verniquet grâce aux travaux de Jeanne Pronteau (Pronteau, 1986) qui a reconstitué toutes les étapes de la construction du plan, depuis l'achat du titre de commissaire-voyer par Verniquet⁴ jusqu'aux dernières gravures des feuilles du plan en 1799. Bien qu'il s'agisse du plus ancien plan que nous utilisons, ces travaux nous offrent la possibilité de connaître en détail la phase, pour nous cruciale, du levé du plan. Il s'agira d'ailleurs de l'unique plan pour lequel nous connaissons avec précision la période de levé.

2.1.1 Levé et structure

Jeanne Pronteau décrit précisément toutes les opérations qui ont mené à la construction de l'atlas de Verniquet. Nous n'en ferons ici qu'une synthèse rapide et certainement incomplète, mais dont l'objectif est de faire apparaître les différentes couches temporelles qui constituent l'atlas final.

La première chose à noter est que la grande majorité des levés topographiques effectuée par Verniquet ne concernent pas la construction du fameux atlas. Son

3. Le nom complet de cet atlas est *Atlas du Plan Général de la Ville de Paris, levé géométriquement par le citoyen Verniquet, rapporté sur une échelle d'une demi ligne pour toise et divisé en 72 planches.*

4. Verniquet prend en effet la suite de Turgot en 1774 après avoir acheté son titre de commissaire voyer à la ville de Paris pour 100.000 livres.

objectif est, à l'origine, la constitution d'un plan de Paris rue par rue. C'est dans ce but qu'il entame en 1775 le relevé topographique de l'ensemble des rues de Paris, dans un premier temps à ses frais, puis, à partir de 1783, officiellement chargé par l'administration royale⁵ Verniquet termine ses plans de rues en 1785 au 1/144^e pour l'ensemble des rues contenues dans les limites de Paris fixées par Louis XV⁶. Notons que ces plans, bien plus détaillés que l'atlas, contiennent, en plus de la forme des rues, les amorces des parcelles, le numéro ainsi que le nom du propriétaire du moment.

La même année, Verniquet demande et obtient de la part du Roi la construction d'un grand plan de Paris, qui aboutira à l'atlas que l'on connaît aujourd'hui. Après avoir effectué les relevés des rues allant jusqu'à l'enceinte des Fermiers Généraux (Jeanne Pronteau donne ainsi 184 rues nouvellement levées jusqu'en 1788 (Pronteau, 1986)) et désormais assisté par plus de 40 personnes, Verniquet commence à former son plan général. Pour cela, il réutilise les levés des rues de Paris qui lui fournissent les mesures des rues puis replace par triangulation ces différentes mesures dans la ville. Finalement, il assemble toutes les rues les unes après les autres pour former la structure complète du plan. Le plan complet de la ville de Paris est ainsi achevé en 1791, mais sa gravure s'étalera jusqu'en 1799.

Structure du plan

L'atlas de Verniquet a été entièrement construit sur la base de relevés topographiques et trigonométriques. À partir de la méridienne de Paris et de sa perpendiculaire se croisant à l'observatoire, Verniquet a effectué une triangulation couvrant toute la ville et permettant d'assembler et de corriger les différents levés des rues. Cette triangulation et la structure complète du plan sont aujourd'hui connus non seulement par les travaux précédemment cités, mais aussi grâce à un plan fourni par Verniquet à la page 70 de son Atlas. Fourni en annexe A.1, il s'agit d'un plan résumant l'ensemble des points fixes choisis comme repère (nommés *points de station*) et les triangles les reliant. Il définit ainsi 67 points de station dont 8 points principaux et 59 points secondaires fixés sur des monuments remarquables (clochers, colonnes, coupes, etc.). Il rajoute enfin 27 points de station aux alentours de Paris, pour améliorer à la fois les vérifications des calculs et la précision des tracés dans les faubourgs. Les 8 points de station principaux sont fournis avec un plan détaillé qui permet de les placer très précisément encore aujourd'hui. Quant aux autres points, si les 59 stations secondaires peuvent encore être retrouvées car placées sur des éléments bâtis (souvent les flèches ou tours d'églises, de couvents et d'autres bâtiments d'importance), la plupart des 27 points des faubourgs ne peuvent pas être placés avec précision (l'un d'eux indique par exemple seulement *Dans la plaine de Grenelle*). Finalement, le plan entier de Verniquet est découpé en 72 feuilles selon un carroyage fixé sur la méridienne et sa perpendiculaire.

5. Jeanne Pronteau indique ainsi que de 1775 à l'été 1783, Verniquet procède à des relevés qui deviendront officiels à la suite de la déclaration royale du 10 avril 1783. Celle-ci fixe des règles entre la largeur de la rue et la hauteur des constructions qui l'entourent. Le Roi charge alors les commissaires voyers, dont Verniquet fait partie, d'effectuer un levé complet des rues de la ville (Pronteau, 1986, p.409).

6. Déclaration du 18 juillet 1724

2.1.2 Contenu du plan

L'atlas de Verniquet est en rupture avec les plans de Paris qui le précèdent dans lesquels le dessin des bâtiments, des cours ou des îlots est souvent très détaillé, tel celui de Jaillot. Son œuvre est à l'inverse extrêmement épurée : seuls les contours des îlots urbains sont dessinés d'un trait d'encre, l'intérieur étant laissé vide. Seul le dessin du relief, c'est notamment le cas de Montmartre, est riche en détails sans doute plus décoratifs qu'indiquant la véritable dénivellation.

Certains bâtiments sont cependant tracés à l'encre et de façon détaillée. Il s'agit en fait de l'ensemble des biens nationaux, qui seront ajoutés par Verniquet pour répondre à la demande de la Convention en 1793 de produire des plans de l'ensemble de ces biens dans le but de les lotir et les vendre (Pronteau, 1986). Cette demande finale aboutira à la gravure, unique, de l'atlas complet entre 1793 et 1799.

2.2 Plan de Maire

Comme nous l'avons évoqué, Paris connaît d'importantes transformations pendant le règne de Napoléon I^e tant en raison des politiques publiques d'urbanisme (démolition du Grand Châtelet, percement de la rue de Rivoli et du canal de l'Ourq, construction d'abattoirs) que des opérations privées de lotissement dans les espaces des différents biens nationaux qui parsèment le centre de la ville.

Le *Plan de la ville de Paris* (ou *La topographie de la Ville de Paris et de ses faubourgs*) dessiné par Nicolas Maire en 1808 est un témoin privilégié des modifications apportées par le Premier Empire, non seulement sur le tissu urbain mais aussi sur les dénominations des rues (on remarque plus de 66 nouveaux noms de rues, auxquelles il faut ajouter les très nombreuses rues ayant repris le nom qu'elles avaient avant la Révolution). Divisé en 20 planches plus une planche d'assemblage, le plan est accompagné par un tableau des rues et une notice explicative de l'auteur qui le présente comme une version corrigée et mise à jour du plan de Delagrive⁷. L'échelle n'est pas notée sur le plan, mais Maire indique qu'elle est la même que celle du plan de Delagrive, soit environ 1/4.500^{ème}⁸, faisant de ce plan l'un des plus détaillés sur les années 1800-1810. À titre de comparaison, une autre série de plans de Paris de haute qualité édités par Charles Piquet à partir de 1804 ne sont qu'à l'échelle de 1/8.000^e.

Nous ne connaissons pas précisément les dates de réalisation de ce plan, 1808 étant seulement sa date de parution. Cependant, il ne s'agit pas d'une édition originale mais de la mise à jour d'un autre plan édité par le même auteur en 1803. Maire corrigera et éditera de nouveau son plan régulièrement, en 1812, 1813, 1816 et finalement 1824 (Bonnardot, 1851, p.226). Un tel rythme de gravure, au moins

7. Précurseur du plan de Verniquet, le plan de De La Grive a été publié en 1728 puis réédité plusieurs fois au cours du XVIII^e siècle, jusqu'en 1761

8. Le plan de Delagrive daté de 1728 est gravé à 11mm pour 25 toises, soit environ 1/4.429^{ème}. Jean Boutier (Boutier, 2007) aboutit également à une échelle de 1/4.500^e, soit 17,2 cm pour 400 toises.

jusqu'en 1816, laisse penser qu'il s'agit surtout de corrections apportées à un tracé réalisé sans doute en 1803 ou avant.

On pourrait s'étonner que l'auteur, pourtant *ingénieur géographe* reconnu⁹, ne se soit pas plutôt appuyé sur l'atlas de Verniquet. Il faut rappeler, cependant, que, tout en restant la référence principale pour la cartographie de la capitale, en 1803 Edmé Verniquet est engagé dans un important procès avec les administrations napoléoniennes et les exemplaires de son Atlas se trouvent dispersés entre les différents protagonistes.

2.2.1 Structure du plan

Un plan issu de plusieurs sources

En premier lieu il nous faut remettre en question la présentation officielle du plan de Maire comme héritier d'un unique plan qui, rappelons-le, est âgé de près de 80 ans. C'est ce que Alain Fourquier suggérait déjà (Fourquier, 2007) en montrant que Maire avait certainement utilisés aussi deux autres plans. Le premier, dit *Plan de Jaillot* (Jaillot, 1775), daté de 1775 à 1778, est proche du plan de Delagrive tant en termes de contenu que de construction : représentation des bâtiments publics et religieux sans détailler les îlots, échelle proche (environ 1/3600^e), etc. Le second n'est autre que l'atlas de Verniquet.

Si les plans de Jaillot et de Delagrive ont pu servir à Maire pour dessiner l'intérieur de Paris, ils ne couvrent pas les espaces externes à l'enceinte des Fermiers Généraux, pourtant représentée dans son plan. Or il est intéressant de remarquer que les espaces de Monceau, Chaillot, Vaugirard et le sud de la Bièvre, en dehors des limites du plan de 1728, étaient parfaitement renseignés chez Verniquet. L'utilisation de son plan par Maire est donc plus que crédible, d'autant plus que les deux géomètres sont contemporain (seuls 4 ans séparent la fin de la gravure de l'atlas de Verniquet et la première publication de Maire en 1803 et que les mises à jours de Verniquet étaient beaucoup plus récentes.

Si l'on compare visuellement les faubourgs de la ville sur les plans de Maire et Verniquet, la ressemblance des styles apparaît immédiatement. La figure 2.1 donne ainsi deux exemples extraits de ces deux plans, le premier centré sur les Buttes-Chaumont à l'Est et le second sur la Butte-Aux-Cailles. On remarque ainsi qu'en plus d'une symbolisation très proche, les mêmes détails sont représentés sur les deux plans. De plus, dans l'espace des Buttes-Chaumont les dénivellations sont presque identiques et symbolisées de la même façon, un tel niveau de détail ne se trouvant que dans l'atlas de Verniquet. À côté de chaque extrait figure également la zone correspondante sur le plan de Delagrive, permettant de mieux juger de la proximité entre les travaux de Verniquet et Maire. Il semble donc que le plan de Maire soit en fait composé de deux sources, la première sur le centre parisien et la seconde permettant la mise à jour des espaces fortement transformés depuis 1728, en particulier autour de l'enceinte des Fermiers Généraux.

9. On le retrouve sous la Restauration avec le titre de *géographe ordinaire du Roi*

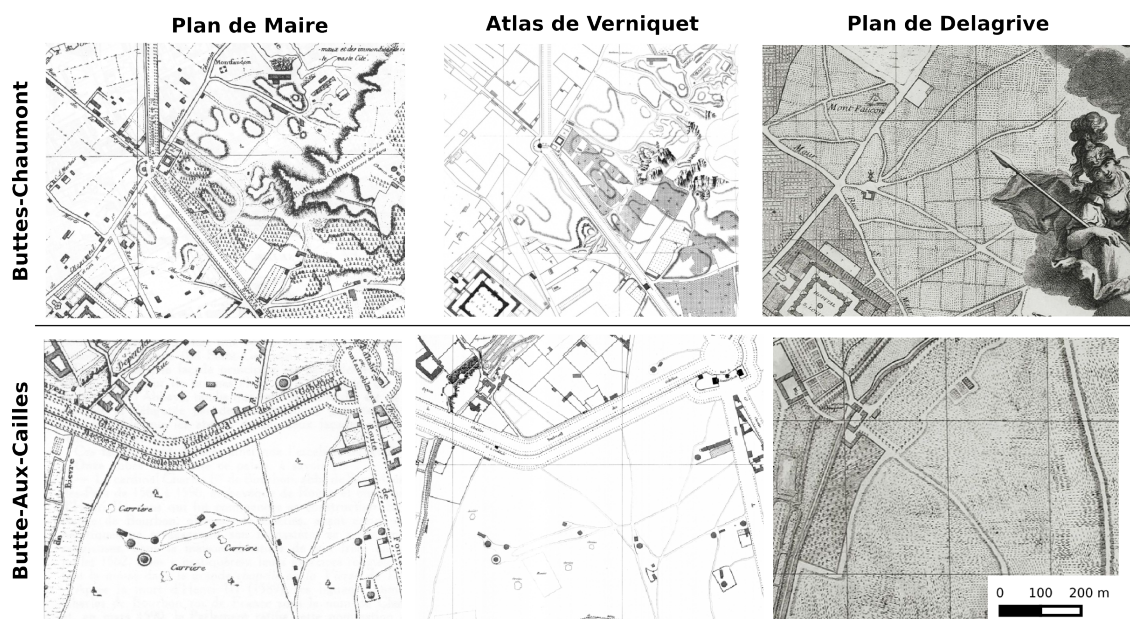


FIGURE 2.1 – Les Buttes-Chaumont et la Butte-Aux-Cailles telles que représentées dans le plan de Maire en comparaison avec l’atlas de Verniquet et le plan de Delagrive.

Le plan de Maire est donc particulier puisqu’il apparaît comme un assemblage d’au moins deux plans plus anciens et non comme une mise à jour d’un plan unique ou un nouveau levé. Bien que cet assemblage ait pu permettre d’obtenir un plan détaillé pour l’ensemble de la ville en 1808, il est donc évident qu’une telle construction peut induire une précision géométrique hétérogène.

Le carroyage du plan

Assez étrangement, Maire ne fait pas apparaître la méridienne de Paris sur son plan, et il ne construit pas non plus son carroyage à la façon de Delagrive et Verniquet en utilisant la méridienne et sa perpendiculaire passant à l’Observatoire pour former les carreaux. Au contraire, le carroyage, composé de 80 rectangles de 900 mètres sur 640 mètres, présente un décalage de plus de 300 mètres à l’ouest de la méridienne, et de près de 150 mètres au nord de sa perpendiculaire. Il corrigera cependant ceci à partir de 1813 en reproduisant le carroyage adopté par Verniquet.

2.2.2 Contenu du plan

Corrections du plan et profondeur historique

Le plan de Maire, comme la plupart des autres plans topographiques de la ville, présente, en superposition du réseau routier, les tracés des voies projetés des travaux de voirie en cours ou prévus. Maire précise que les tracés annotés sur le plan sont tous arrêtés par le gouvernement, ce qui n’empêche pas le plan de contenir certains projets jamais réalisés. Parmi les travaux principaux indiqués, le pont Iéna en 1814, le prolongement de la rue de Tournon ou encore la rue d’Ulm verront le jour. Le canal de l’Ourq est quant à lui noté comme projet et esquissé au sud de son tracé. Les réorganisations des abords de l’Observatoire de Paris et de l’église Saint-Eustache,

jamais effectuées, sont tout de même décrites.

Sur les rues toujours, Maire fait le choix de figurer les noms des rues révolutionnaires ayant changé de nom pendant le Premier Empire. Puisque le plan de 1808 est une réédition corrigée d'un plan de 1803, ces doubles dénominations concernent les rues ayant changé de nom entre 1803 et 1808. Par exemple, la rue du Roi de Sicile qui s'est appelée *rue des droits de l'Homme* pendant la Révolution est visiblement corrigée sur le plan qui indique *Rue du Roi de Sicile et non Rue des Droits de l'Homme*. Il s'agit donc davantage de corrections de l'édition précédente que d'une volonté de description historique. De la même façon, la maison de chanoine du Petit Saint Antoine qui se situerait aujourd'hui au 11 rue du Roi de Sicile, détruite en 1792 et définitivement démolie en 1804 est dessinée sur le plan et accompagnée d'une note précisant sa destruction : *démoli en 1804*. Maire ne se contente pas de corriger son plan mais prévoit certains changements de noms par avance, telle la rue de la Réunion, ancienne rue de Montmorency, qui reprendra son titre d'origine en 1808. Aussi Maire note-t-il : *rue de Montmorency act. rue de la Réunion*. D'autres changements plus tardifs sont indiqués de la même façon, comme la rue révolutionnaire *De la Convention* qui deviendra en 1814 la *rue du Dauphin* ou encore l'ancienne place du Louvre appelée place d'Iéna jusqu'en 1814 que Maire renomme *Place de la Colonnade*.

Dernière particularité de ce plan, plusieurs planches sont annotées de textes décrivant certains bâtiments de la capitale impériale ou certains lieux qui en offrent un point de vue unique. Ainsi, la planche 10 du plan, centrée sur le Champ de Mars, indique qu'*ici l'École Militaire se déploie dans toute sa beauté, au milieu de la campagne la mieux cultivée avec son vaste et magnifique Champ de Mars*. Le plan de Maire se veut précis et juste mais vise aussi un usage touristique dont l'objectif est d'intéresser *l'Étranger comme l'habitant de Paris* pour lui donner un accès simple à *tout ce que [la] capitale de l'Empire renferme d'utile et de curieux*. (Maire, 1808, p.VII). Le plan de Maire est l'unique plan touristique dont nous ferons usage au sein de ce travail, ceux-ci étant souvent très peu précis géométriquement et, par là même, difficilement utilisables au sein d'un SIG. Le plan de Paris réalisé par Nicolas Maire est la seule de nos sources de données qui ne soit pas directement liée aux dépositaires officiels de l'aménagement urbain parisien. En effet, jusqu'à la réalisation du cadastre parisien en 1807, le travail de Verniquet constituait l'unique grand plan utilisé par l'Hôtel de Ville, qui ne possédait alors aucun service cartographique suffisamment développé pour produire un nouveau grand plan de Paris. De 1791 à 1810 au moins, nous sommes donc obligés de nous fier à des plans réalisés par des cartographes indépendants, dont l'exactitude et la fiabilité sont souvent très critiquables. Choisir le plan de Maire implique donc de s'intéresser à un plan singulier par sa construction mais aussi particulièrement riche de contenu. L'ultime édition date de 1824 et continue à intégrer des transformations en présentant les plaines d'Ivry et de Châtillon, lieux de plaisir et de guinguettes où se rendaient tant de Parisiens. Ces plans étaient vendus en noir et blanc, ou coloriés.

2.3 Cadastre de Vasserot

Réalisé dans un premier temps par l'architecte employé des services du cadastre Philibert Vasserot, puis par son fils Charles Vasserot, le cadastre *par îlots* de Paris est une production cartographique majeure du XIX^e siècle. Celui-ci se présente sous la forme de 912 plans des îlots de la ville, rassemblés en 24 volumes (Noizet, 2009; Pinon *et al.*, 2004). La quasi totalité des 12 anciens arrondissements est couverte. Nous avons utilisé pour notre travail les plans d'îlots assemblés et géoréférencés dans le cadre du projet ALPAGE (Noizet *et al.*, 2008). Ce projet ayant mené à des recherches sur la création de l'atlas lui-même et sur le personnage de Philibert Vasserot, nous nous appuyons ici sur les travaux de Noizet (2009). On pourra également trouver une étude détaillée du plan et de son auteur dans Souchon (2002) et des détails complémentaires dans Coyecque (1909).

Une part du cadastre napoléonien

Après l'échec du premier cadastre français en 1802 *par masses de culture*, Napoléon Bonaparte relance avec la loi du 15 septembre 1807 le projet d'un cadastre parcellaire sur la France entière, immédiatement mis en œuvre. La capitale, qui en 1802 n'avait pas été couverte par les opérations, connaît une fois encore un traitement particulier. En effet, l'Atlas de Verniquet, alors utilisé par les services de la ville, constitue un plan suffisamment précis pour être utilisé comme base d'un plan cadastral. Son trait épuré le rend particulièrement adapté à servir de fond topographique pour ce travail. C'est donc assez logiquement qu'en 1807 le Service du Cadastre de la préfecture entame la réalisation d'un plan parcellaire des 48 quartiers de Paris au 1/2.000^e fondé sur l'Atlas de Verniquet, plan qui sera terminé en 1821. Mais l'échelle de ce plan ne permet pas de décrire en détail une ville aussi dense et il se contentera de représenter les amorces des parcelles avec leurs numéros. En 1809, un second projet cartographique est lancé pour compléter ce premier travail en se plaçant à une plus grande échelle. Cette fois, il s'agit de lever le plan de toutes les maisons de la capitale¹⁰, ce qui prendra 27 ans et se terminera en 1836. Philibert Vasserot, architecte des Hospices de Paris et attaché au service du cadastre, travaillera sur ces deux projets.

Constatant que l'écart d'échelle entre les deux plans complique la localisation d'un plan de maison dans celui des 48 quartiers, Philibert Vasserot propose alors la réalisation d'un plan intermédiaire par îlots devant former *un atlas composé de 60 à 80 tableaux par arrondissement municipal*. Cécile Souchon (Souchon, 2002, p. 3) précise que le travail de Vasserot, commencé en 1810 avancera à bon rythme jusqu'en 1815 avant de ralentir et finalement de s'arrêter faute de financement dans les années suivantes. Il reprendra son travail en 1821, qui accélérera à partir de 1823 quant Philibert Vasserot recevra l'aide d'autres personnes¹¹ et sera publié en

10. On trouvera une note détaillant précisément les opérations de levé des maisons dans (Souchon, 2002, p. 2)

11. Souchon (2002); Coyecque (1909); Noizet (2009) s'accordent sur l'aide apportée par son fils Charles Vasserot, qui poursuivra d'ailleurs le travail de son père. Noizet (2009) s'interroge sur la participation d'autres personnes à partir de 1823, notamment un architecte du nom de Delucenay qui prendra le relais de Charles Vasserot après 1848.

1836 bien qu'il soit alors inachevé. Philibert Vasserot meurt en 1840, mais son fils poursuivra le dessin des îlots manquants jusqu'en 1848 avant de laisser lui même la main à l'architecte Delucenay. Noizet (2009) fait en outre remarquer que l'atlas régulièrement appelé *atlas de Vasserot et Bellanger* est en réalité un projet initié par Vasserot après 1836 visant à rassembler les plans d'îlots par quartiers. Le projet ne sera toutefois jamais terminé et seules les minutes¹² de 37 quartiers tracées au crayon seront faites.

2.3.1 Levé et structure du plan

Les plans des îlots levés par Vasserot sont tout à fait particuliers. Contrairement à la majorité des plans topographiques qui sont construits selon une logique *descendante*, partant d'une triangulation générale de l'espace qui est ensuite raffinée, Vasserot s'appuiera directement sur les plans des maisons. Nous ne savons pas si Vasserot s'est appuyé sur les plans d'îlots déjà réalisés par Verniquet entre 1785 et 1791 (Noizet, 2009, p. 5), peut-être également utilisé par les géomètres ayant réalisé le plan des 48 quartiers, ou s'il a effectué de nouveaux levés. Il est toutefois certain que les plans d'îlots ont été réalisés en assemblant la multitude de plans de maisons levés en même temps que les feuilles étaient dessinées.

Une telle méthodologie pose deux problèmes. D'abord, la réalisation simultanée des îlots et des plans de maisons contraignait Vasserot à attendre la réalisation des maisons d'un îlot pour pouvoir le terminer. Or, il apparaît que les levés des maisons des deux côtés d'une rue étaient parfois décalés de plusieurs années, menant éventuellement à des plans représentant deux états différents de la rue. Plus encore que les autres plans, le cadastre par îlots de Vasserot contient de multiples temporalités. Le projet ALPAGE a analysé l'ensemble des 912 îlots référencés et a déduit pour chacun d'entre eux une date approximative de dessin.

Ensuite, le fait d'assembler des plans de maisons réalisés indépendamment questionne la précision géométrique du résultat. En effet, la juxtaposition de ces plans inévitablement inexacts mène à une accumulation d'erreurs aboutissant à des îlots dont les limites sont distendues. Si en revanche les limites des îlots étaient connues de Vasserot par de nouveaux relevés ou grâce à Verniquet, les plans de maisons ont pu être modifiés pour pouvoir s'agencer à l'intérieur. Dans tous les cas, la création de plans en assemblant d'autres plans demande à s'interroger sur leur précision géométrique et peut engendrer des déformations potentiellement importantes lors du géoréférencement.

2.3.2 Contenu du plan

De par son échelle et l'ampleur de la tâche (plus de 40 personnes -ingénieurs géomètres, architectes et arpenteurs- participeront au levé des plans des maisons), le cadastre de Vasserot est extrêmement détaillé. De plus, contrairement aux autres

12. La minute désigne les premières versions du plan, généralement dessinées sur le terrain, qui précèdent la gravure ou bien la mise au net.

plans que nous utilisons, construits comme des outils d'aménagement et de promotion de Paris, le cadastre de Vasserot a un but fiscal. L'accent est donc mis sur la parcelle plutôt que sur la voirie ou la ville elle-même. Le réseau des rues n'est visible que dans les plans d'assemblage de chaque quartier fourni avec l'atlas par îlots, mais n'est pratiquement pas représenté dans les plans d'îlots, dont la nature même ne permet pas de représenter une rue dans son intégralité.

À l'inverse, les propriétés foncières sont extrêmement détaillées : toutes les maisons sont vues en coupe, présentant leur configuration interne, mais aussi les puits, cours, escaliers, etc. Noizet (2009) note cependant que certains éléments distingués sur les plans n'ont sans doute pas fait l'objet de mesures spéciales et sont ainsi représentés uniquement dans un but fiscal ou pour faciliter la lecture. C'est par exemple le cas des murs internes et externes des maisons (les premiers étant plus épais sur le plan), ou encore des escaliers internes aux maisons.

Enfin, les plans indiquent deux numérotations des maisons. Les premiers, notés à l'encre, sont les numéros impériaux de 1805. D'autres numéros, indiqués au crayon, sont des corrections et mises à jour effectuées sans doute par le personnel du cadastre.

Les plans par îlots de Vasserot, bien que jamais terminés, ont été utilisés par la préfecture de la Seine comme plan parcellaire, comme en témoignent divers plans d'alignements de rues¹³ mais aussi la mise à jour des numéros d'adresses rues qui y figurent.

Pourtant, les services de la Seine étaient conscients dès 1835 que la durée de réalisation des plans constituait un sérieux frein à leur utilisation. Accumulant des plans faits à des périodes trop différentes, dessinant une ville en pleine transformation, ils devenaient rapidement obsolètes ou incohérents et obligeaient à rapporter en continu les nouvelles rues et maisons. Comme le constate Souchon (2002), le projet de Philibert Vasserot était en fait voué à l'échec, tant les temps de réalisation des feuilles étaient inadaptés au rythme de transformation d'une ville en pleine expansion.

2.4 Atlas de Jacobet

Simon-Théodore Jacobet, né en 1798 à Toulouse fut architecte employé dès 1823 à la Préfecture de la Seine puis chef du bureau chargé de la réalisation des plans d'alignements. Mêlé à divers procès liés à ses activités à la préfecture, il fut révolutionnaire en 1830, 1832 puis 1848, arrêté, interné et condamné à la déportation en Algérie en 1852, condamné à la *mort civile* et enfin assigné à résidence à Montesquieu-Volvestre la même année. Il sera l'auteur du plus grand et plus complet plan de Paris existant sur la première moitié du XIX^e siècle.

La réalisation de son *Atlas Général de la Ville, des faubourgs et des monuments de Paris* est une fenêtre ouverte non seulement sur la topographie parisienne pré-haussmanienne, mais aussi sur le fonctionnement des services de voirie de la Seine.

13. En 1851 encore, les plans de percements de la rue de Rivoli entre la rue de la Bibliothèque et la rue du Louvre seront tracés sur un plan parcellaire très proche de celui de Vasserot

Suivre la construction de cet atlas permet non seulement d'entrer au cœur de la machine d'aménagement urbain mise en place en 1800 par Napoléon, mais surtout de découvrir les rapports qu'entretenaient les employés de la préfecture et les agents d'affaires dans un objectif commun de spéculations immobilières, ainsi que la corruption à l'œuvre dans les services chargés de l'aménagement urbain : percements, alignements, gestion des carrières, etc.

2.4.1 Levé et structure du plan

Le travail de Jacoubet s'inscrit volontairement dans la lignée des grands plans de Paris contruits au XVIII^e siècle¹⁴. Commencé entre 1825 et 1827, l'Atlas Général s'inspire directement des travaux de Verniquet dont il reprend en partie la triangulation. Il est important de noter dès maintenant que Jacoubet entreprend la réalisation de son atlas alors qu'il se trouve employé au *Bureau des Plans* de la préfecture de la Seine. C'est grâce à cette position qu'il sera en mesure d'utiliser les relevés topographiques réalisés par les géomètres de l'administration à partir des plans de Verniquet, alors utilisé comme plan général pour les travaux de voirie. Comme nous le verrons plus tard, il a également repris les mesures de triangulation effectuées par les équipes de Verniquet, ce qui lui permet d'économiser des opérations de levé topographiques d'envergure¹⁵. L'atlas est réalisé entre 1825 (ou 1827) et 1836 et il est publié par parties selon la méthode de la souscription. Cette méthode consiste à financer le travail de gravure, très coûteux, par étapes successives grâce à la contribution de particuliers qui subventionnent des lots de feuilles d'atlas. Au total, *l'Atlas Général de la ville de Paris* comporte 54 feuilles traçant un plan de Paris au 1/2000^e¹⁶. Les feuilles 53 et 54 présentent en outre un plan des principales opérations de triangulation ayant permis de construire l'atlas.

2.4.2 Contenu du plan

Tout comme l'atlas de Verniquet, Jacoubet crée un plan relativement épuré contenant principalement le tracé des rues et les plans des bâtiments publics. Toutefois, l'objectif de l'architecte est de faire de son atlas un outil de travail pour la préfecture de la Seine, mais aussi pour les propriétaires et entrepreneurs parisiens. Pour cette raison, il rajoute le tracé des alignements prévus par la préfecture en vertu de la loi du 16 septembre 1807 (cf. le paragraphe 2.6.1), ainsi que les parcelles cadastrales numérotées. Enfin, les bâtiments à l'intérieur des boulevard sont dessinés en coupe ; On peut d'ailleurs remarquer, par une étude fine des planches et de l'espace qu'elles représentent, que les échelles des bâtiments et des autres thèmes cartographiques ne sont pas toujours identiques. L'échelle des bâtiments est ainsi régulièrement plus grande que celle des rues et îlots. Cela s'explique par le fait que les bâtiments figurés dans l'atlas proviennent très certainement des levés de l'atlas des 48 quartiers de

14. Notamment ceux de Delagrive, Verniquet, Delisle et Jaillot

15. Paris ayant toutefois évolué depuis 1791, notamment aux alentours des boulevards, Jacoubet complète le canevas de triangle existant Réutilisant en partie les levés de Verniquet, il est possible qu'il se soit cantonné à lever en détail les parties périphériques de la ville. En effet, contrairement à Verniquet ou même Vasserot, Jacoubet est presque seul à réaliser son atlas. Seuls quelques employés de la Seine l'aideront à reporter les calques -c'est à dire les premières minutes- de l'atlas sur les feuilles de l'atlas.

16. L'échelle indiquée sur le plan est de *1 millimètre pour deux mètres*

Vasserot, différents de ceux de Verniquet. On a ici un exemple de réutilisation de différentes sources cartographiques générant des erreurs dans la carte ainsi consituée en *patchwork*. L'atlas est donc globalement hétérogène. Tout d'abord, les bâtiments à l'extérieur des boulevards sont dessinés en masse, à l'inverse de paris intra-muros. Le dessin des parcelles est également très inégal. Dans l'extrême centre de Paris (autour de la place du Châtelet) et de l'extérieur de la ville, toutes les parcelles sont représentées et numérotées. Partout ailleurs, les parcelles sont seulement ébauchées et seuls leur numéro et leur amorce sur la rue est dessinés.

Tous ces éléments contribue à faire de l'atlas de Jacoubet un plan majeur du milieu du 19^e siècle mais dont les hétérogénéités appellent à le considérer avec prudence. Cet atlas et son auteur sont symptomatiques de la mutation que subit la gestion urbaine au 19^e siècle, entamée entre la Révolution et le Premier Empire et qui s'achèvera par l'arrivée du préfet Haussmann à la tête de la très centralisée préfecture de la Seine. Pour cette raison, nous proposons en section 2.6 d'explorer plus en profondeur le personnage de Jacoubet et la réalisation de son grand atlas, ce qui nous permet de mettre en évidence cette mutation.

2.5 Atlas municipal

Nous avons vu précédemment que les plans de Vasserot ont été utilisés au moins jusqu'en 1850, peut-être jusqu'à la réalisation d'un second cadastre sous l'impulsion de Haussmann. Les plans de Vasserot ne sont cependant pas des vues d'ensemble de la capitale et sont inadaptés pour la mise en place de projets de voirie globaux. Nous ne savons cependant pas quels plans ont rempli ce rôle de 1843 à 1854, si ce n'est l'énigmatique plan entamé sous Frochot et relancé en 1842 sous l'impulsion de Planson. L'atlas de Jacoubet n'a quant à lui sans doute pas été utilisé à grande échelle par la préfecture, les droits d'exploitation et les cuivres restant entièrement aux mains de son auteur.

L'arrivée de Haussmann à la tête de la préfecture de la Seine en 1853 bouleversera le statut des plans topographiques de la ville, bien plus que ne l'avait fait la réorganisation du Service des Plans en 1843. En effet, chargé de mettre en œuvre le programme d'urbanisme élaboré depuis plusieurs années par l'empereur Napoléon III, le préfet Haussmann aura pour priorité la création de deux grands plans de Paris destinés à planifier et suivre les travaux de voirie aujourd'hui bien connus. Ainsi après avoir totalement réorganisé et agrandi le Service des Plans en 1856, renommé pour l'occasion Service du Plan de Paris, il place à sa tête l'ancien conservateur Eugène Deschamps (voir la figure 2.5). Haussmann ordonne alors le levé d'un nouveau plan topographique de Paris ainsi qu'un nivellement complet de la capitale. Nous nous intéresserons ici seulement au plan topographique au 1/5.000^e qui sera réalisé à partir de 1856 sur la base d'une triangulation entièrement nouvelle effectuée entre 1856 et 1857, terminé en 1862 et publié pour la première fois en 1864. D'abord appelé *Atlas administratif* puis à partir de 1868 *Atlas municipal*, le grand plan de Haussmann sera publié à un rythme quasi-annuel jusqu'au XX^e siècle¹⁷ sous

17. Sont portées à notre connaissances les éditions de 1864, 1866, 1868, 1876, 1878, 1880, 1882, 1885, 1886, 1887, 1888, 1893, 1894, 1895, 1898

la forme d'un atlas de 21 feuilles d'abord, puis 16 à partir de 1866. Il existe en outre des éditions du même atlas ramené à l'échelle du 1/10.000^e.

Chaque édition est scrupuleusement mise à jour par le service du plan et représente la ville l'année de sa publication. Pour notre part, nous nous sommes servis uniquement de l'édition de 1888 intitulée *Atlas municipal des vingt arrondissements de la ville de Paris. Dressé sous la direction de M. Alphand inspecteur général des ponts et chaussées, par les soins de M.L Fauve, géomètre en chef, avec le concours des géomètres du plan de Paris* (Alphand et Fauve, 1888) réalisé sous la direction du préfet Eugène Poubelle.

2.5.1 Structure du plan

Triangulation

L'atlas municipal est la version publiée et compilée dans un ouvrage du grand plan réalisé sous la direction de Haussmann en 1856. Ce plan, également au 1/5.000^e a été dessiné vers 1857 en deux exemplaires. Le premier, placé sur un grand tableau de bois, était disposé derrière le bureau du préfet et servait à reporter toutes les opérations de voirie réalisées ou projetées. Ce plan semble avoir aujourd'hui disparu, mais il a peut être survécu à l'incendie de l'Hôtel de Ville. En effet, on en trouve une dernière trace dans un rapport de Edgar Mareuse à la commission du Vieux Paris en 1925 (Mareuse, 1925). Le second, parfois appelé aujourd'hui *plan de Morizet* du nom de l'historien l'ayant redécouvert dans les années 1930, représentait tous les projets prévus par Napoléon III et avait été offert au roi de Prusse Guillaume I^{er} en 1867.

Pour tracer ce plan, Haussmann indique dans ses mémoires (Haussmann, 1893) qu'une nouvelle triangulation complète de Paris a été effectuée entre 1856 et 1857, sous la direction d'Eugène Deschamps¹⁸. Ainsi, ce plan constituerait la première triangulation complète effectuée depuis l'atlas de Verniquet, du moins si l'on conserve l'hypothèse d'une triangulation seulement partielle pour Jacoubet. Haussmann, affirmant qu'aucun grand plan de Paris n'existait lors de son arrivée à la préfecture renouvelle ainsi totalement les outils de l'administration. Il n'est cependant pas certain que les choses aient été si simples et la tendance du préfet à se placer en fondateur de la cartographie officielle en omettant volontairement des projets antérieurs a déjà été pointée par Pierre Casselle (Casselle, 2000) (C'était alors la commission des embellissements du Comte Siméon qui était ignorée et son rôle auprès de l'empereur minimisé). En effet, le frontispice d'une édition réduite au 1/10.000^e conservée à la Bibliothèque Nationale de France (Deschamps *et al.*, 1871) indique que *le grand plan en 21 feuilles dressé à l'échelle de 1/5000 résume les travaux des géomètres du Plan*

18. Il s'agit d'une opération d'une ampleur inédite. Dans tout Paris, des mâts de bois munis d'une plateforme dépassant la hauteur des maisons sont plantés sur les points de triangulation définis dans les rues, sur les places, etc. Les géomètres peuvent alors monter au sommet pour mesurer les angles aux autres mâts et bâtiments et effectuer de nombreuses vérifications, tandis que des arpenteurs mesurent les distances entre les mâts (Haussmann, 1893, p. 14)

de Paris depuis 1842.¹⁹. Cette date coïncidant avec la reprise des travaux de levé et de dessin sur le plan en 48 parties de Frochot, on peut alors se demander si ce ne sont pas de ces travaux dont parle cette note, qui, ajoutés aux nombreux plans levés par les services de la voirie entre 1842 et 1856 ont pu servir aux géomètres sinon à dessiner le nouveau plan de Paris, du moins à guider les opérations de triangulation en aidant à définir les emplacements et le nombre de points de station nécessaires.

Structure

L'édition que nous utilisons est composée de 16 feuilles carroyées, couvrant l'intégralité du Paris de 1888, soit l'ancienne ville (décrite par les plans précédents) ainsi que les communes annexées en 1860. Le carroyage, fixé sur la méridienne de Paris et sa perpendiculaire forme des carreaux de 1.000 mètres de côté. Au même titre que les atlas de Verniquet et de Jacoubet, l'atlas de Haussmann vise la plus parfaite précision géométrique.

2.5.2 Contenu

Édité sous la direction d'Adolphe Alphand qui avait réalisé de nombreux jardins, parcs et promenades sous la direction de Haussmann, l'atlas de 1888 donne des représentations très précises des différents aménagements paysagers de la ville, en particulier des bois de Vincennes et Boulogne dont des plans individuels avaient été dressés dans les années 1860. C'est aussi le premier plan que nous utilisons qui figure les trottoirs et les chemins de fer de la ville, bien que ceux d'Austerlitz, d'Orsay et du Nord soient présents sur les éditions plus tardives de l'atlas de Jacoubet. Aucune parcelle n'est visible, l'atlas abandonnant les représentations hybrides figurant seulement l'amorce des parcelles, ici inutiles puisque la préfecture s'était doté également d'un nouveau plan parcellaire. Les numéros des maisons sont indiqués seulement à la fin des tronçons, et sur les bâtiments publics. Enfin, aucune habitation particulière n'est représentée, mais tous les bâtiments public sont dessinés et nommés.

Le plan de Haussmann est donc le premier plan véritablement officiel de Paris, créé par l'administration de la Seine dans l'unique but de servir d'outil à l'aménagement de la ville. Il ne s'agit plus, comme pour l'atlas de Jacoubet, d'un produit destiné également à être vendu au grand public. Les versions publiées du plan de Haussmann représentent la ville uniquement sous l'angle de l'administration, seuls les éléments officiels de la ville étant dessinés.

Reste enfin la question de la filiation du plan de Haussmann. Nous avons vu que subsistaient quelques doutes sur le fait que ce plan soit à l'initiative unique de Haussmann. Il a cependant eu la volonté, la puissance financière et le pouvoir suffisants pour enfin dresser un plan qui était en projet au moins depuis 1842. Il s'agit

19. Le cartouche complet indique : *Plan de Paris, réduit à l'échelle de 1/10000 d'après le grand plan en 21 feuilles. Le grand plan en 21 feuilles dressé à l'échelle de 1/5000 résume les travaux des géomètres du Plan de Paris depuis 1842. Il a été exécuté sous la direction de M. Eug. Deschamps chef du Service du plan de Paris, par les soins de MM. de Lucenay Ross et Chevigny géomètres triangulateurs Berger Montry Picard-Dobré et Pozier géomètres en chef des brigades topographiques et L. Fauve, géomètre en chef du Service intérieur chargé de la réduction. Il a été gravé sur pierre par MM. Avril frères entre 1866 et 1871.*

d'une brisure dans l'utilisation des plans de Paris par la préfecture, faisant passer définitivement les grands plans de Paris dans le giron de la préfecture. D'ailleurs, l'atlas de Jacoubet sera le dernier plan détaillé de Paris d'initiative privée.

2.6 L'atlas de Jacoubet : des architectes voyers à l'administration de la Seine

Nous venons de présenter les cinq plans que nous allons utiliser pour construire la base de notre système d'information géographique. Notre choix était pratiquement obligé. Comme nous venons de dire, ils constituent la quasi totalité des bases utilisées comme référence pour la production des cartes parisiennes depuis la révolution à la fin du XIXe siècle. Or, malgré cette présentation schématique, nous avons aussi essayé de montrer comment, au delà d'une indéniable continuité formelle, d'ailleurs donnée par la nature même de l'objet représenté, chaque plan naît et se développe dans une configuration politique, technique et administrative, qui évolue incessamment au cours de cette période tourmentée. D'une part le plan le plus ancien, conçu avant la révolution et fini de graver sous le Directoire, est produit par l'architecte privé Edmé Verniquet à destination d'une administration publique. D'autre part, l'Atlas municipal, est une œuvre conçue et développée totalement à l'intérieur de la nouvelle administration centrale mise en place par le préfet Haussmann. Ces deux plans constituent donc les points extrêmes du processus qui voit la centralisation progressive des services administratifs de la préfecture et de la ville de Paris et qui s'accompagne à la naissance de la figure du fonctionnaire technicien. Si l'analyse de ces dynamiques pourrait sembler secondaire, voir même inutile, à l'intérieur d'une étude dédiée à la construction d'un modèle de SIG historique, elle apparaît indispensable pour saisir correctement la nature et la validité des sources cartographiques utilisées. Car c'est bien dans la tension entre fonction publique et fonction privée, ou encore entre les emprunts et les réutilisations des uns sur les autres, que se produisent les représentations d'un espace qui change, lui aussi, d'ailleurs, de manière très conflictuelle.

C'est donc dans cet horizon qu'il est indispensable de se pencher de manière plus approfondie sur la figure et l'œuvre d'un personnage comme Théodore Jacoubet qui condense, dans son parcours, l'ensemble de ces tensions et de ces dynamiques. Architecte privé, employé de la Seine, savant reconnu et farouche républicain, Jacoubet entame sa carrière au moment où, avec la construction du cadastre, se pose la nécessité de construire un plan général de la ville en mesure de substituer les anciens levés de Verniquet sur lesquels s'était calé Vasserot. Nous sommes au début de la Restauration et dans les 30 ans qui se sont écoulés depuis la révolution, des transformations majeures ont affecté la ville de Paris : lotissements de grande ampleur, premiers grands percements, alignements généralisés... Les plans de Verniquet et même celui de Vasserot sont rapidement obsolètes. Pourtant, entre 1821 et 1854, on ne connaît pas de plan général de la ville, alors même que Rambuteau s'affiche comme *préfet voyer*. Théodore Jacoubet, son atlas et ses rapports avec la préfecture vont nous permettre de voir, au moins de loin, l'intérieur des bureaux de la préfecture et par là nous donner quelques éléments pour compléter l'histoire des

grands plans de Paris pendant ce singulier vide (bien que partiel, n'oublions pas que les plans de Vasserot seront en production jusqu'en 1854).

2.6.1 De la servitude d'alignement à l'Atlas Général

La loi du 16 septembre 1807 et ses effets

Avec le cadastre de Vasserot, nous avons vu comment s'est concrétisée, pour la ville de Paris, la traduction cartographique de la loi du 15 septembre 1807. Celle-ci est certainement la plus connue de la période puisqu'elle reprend et structure le cadre juridique d'un outil administratif dont le principe reste aujourd'hui relativement inchangé. Pourtant, une seconde loi, datée du 16 septembre 1807²⁰ est, à l'échelle de Paris au moins, d'une importance capitale. Ses effets sur la cartographie de la ville sont encore moins connus car ils ne se manifesteront que très graduellement au cours des vingt années suivantes. Reprenant les déclarations royales de 1783 portant sur la largeur des rues de Paris²¹, cette loi fixe le cadre d'un outil d'aménagement urbain toujours existant : la servitude d'alignement, qui permet aux maires (à Paris à la préfecture) de dresser des plans d'alignements des rues fixant les limites que les maisons ne peuvent pas dépasser, garantissant ainsi une largeur minimale des voies.

Elle complète également la législation sur l'expropriation pour cause d'utilité publique, dotant finalement les administrateurs de trois outils d'aménagement :

- L'alignement, qui consiste à définir des limites d'avancement que les maisons ne peuvent pas franchir, qui s'applique surtout aux nouvelles constructions.
- Le reculement, lequel consiste à déplacer la façade d'un bâtiment pour qu'il respecte la largeur de rue officielle. Le reculement peut être fait spontanément par le propriétaire, mais la plupart du temps des accords sont passés au cas par cas entre les individus et la préfecture. La préfecture peut également appliquer des amendes aux contrevenants, les obligeant par là à reculer d'eux-mêmes leurs maisons, ou à passer un accord avec l'Hôtel de Ville.
- L'expropriation, où la préfecture saisit tout ou partie d'une parcelle et indemnise le propriétaire en fonction de la surface enlevée et de l'ampleur des destructions. Là aussi les accords entre propriétaires et préfecture sont courants, celle-ci pouvant notamment autoriser la revente des matériaux issus de la démolition.

Tout comme le cadastre parcellaire qui était conçu par Napoléon pour être un outil et un appui du code civil, le plan général d'alignement accompagnera la loi du 16 septembre. Elle impose à toutes les villes de plus de 2000 habitants la mise en place d'un plan général d'alignement des rues, donnant pour chacune les lignes à leurs bords que les constructions ne peuvent dépasser sous peine d'amende et de démolition.

20. Article 52 de la loi du 16 septembre 1807. *[Les] alignements sont également donnés par les Maires ; mais il faut qu'ils soient conformes aux plans qui doivent être arrêtés préalablement au conseil d'état, sur avis du préfet. Ainsi, dans les villes où il y aurait des routes nouvelles à ouvrir, ou d'anciennes à élargir, les Maires doivent faire dresser des plans d'alignements à déterminer, et envoyer ces plans au préfet avec leurs propositions motivées.*(Merlin, 1827). Rappelons que Paris étant privée de maire de 1794 à 1848, c'est la préfecture de la Seine qui est alors directement chargée des alignements.

21. Cf. la déclaration royale du 10 avril 1783, dans (Jousselin, 1850).

Malheureusement à Paris cette loi s'avère pratiquement inefficace en raison de son caractère passif. En effet, l'expropriation pour cause d'utilité publique, pourtant partie de la loi du 16 septembre, n'est que très peu utilisée par la préfecture tandis que la servitude d'alignement est appliquée à toute la ville. Chabrol en donnera trois raisons : de nombreuses rues sont suffisamment spacieuses pour ne pas nécessiter l'intervention directe de la préfecture, d'autres sont de trop faible importance pour s'en occuper et, enfin, pour éviter les opérations de spéculation immobilière (De Chabrol de Volvic, 1823, pp. 197-198). Sans expropriation, les maisons existantes ne respectant pas la largeur de rue prévue ne peuvent être détruites par la municipalité, alors que les nouvelles constructions sont érigées avec des écarts de profondeurs allant jusqu'à plusieurs mètres. Quant au reculement, s'il est utilisé, les négociations avec les propriétaires des maisons visées sont si longues que le préfet Chabrol prévoira en 1819 que plus d'une centaine d'années seraient nécessaires pour mettre en œuvre tous les alignements prévus. Pire encore, au lieu d'améliorer l'état des rues, la loi participe à sa dégradation. Les recoins entre deux maisons avancées différemment se remplissent d'ordures et d'eau stagnante. Adoptée dans l'optique d'apporter *air et lumière* aux rues de Paris, la loi sur les alignements produira dans Paris les effets inverses. Le comte de Rambuteau, qui prendra la tête de la préfecture de la Seine en juin 1833, mettra en place une politique d'aménagement urbain bien plus agressive que ses prédécesseurs, recourant plus largement à l'expropriation pour faire appliquer le plan d'alignement de Paris.

La mise en place des outils qui permettront les percées et les alignements de la première moitié du XIX^e siècle, en particulier des plans d'alignements des rues et d'un plan général de Paris, sera cependant commencée avant Rambuteau, sous la direction du préfet Chabrol.

1820-1830 : Le projet de plan général de Paris du préfet Chabrol

Dans les premières années du 19^e siècle, Paris ne possède aucun plan cadastral global. Seuls quelques quartiers sont couverts par une première tentative réalisée en 1802 (Pinon *et al.*, 2004, p. 86). Le manque de documents cartographiques pour l'administration étant général pour toutes les villes françaises, le ministère de l'intérieur publie en 1808 un décret donnant aux maires deux ans pour réaliser les plans d'alignements des rues²². Dorénavant, le dessin de plans d'alignements est un préalable obligatoire à toute opération de voirie destinée à élargir, percer ou simplement maintenir l'état d'une rue. Paris n'échappe pas à la règle, mais la levée de plans d'alignements est, dans une ville de cette taille, une tâche de longue haleine (la superficie des terrains à retrancher est estimée à 506.000m²). Pour comparaison, le plan d'alignement de Poitiers (21.000 habitants environ en 1820, contre plus de 700.000 pour Paris à la même époque) occupera plusieurs personnes (un architecte, deux dessinateurs, au moins 2 arpenteurs) pendant 2 ans et demi (Pinon, 1996). Le renvoi du préfet Frochot en 1812 et les troubles de la fin de l'Empire n'ont sans

22. Décret du 27 juillet 1808, article 1^{er} : *Les alignemens qui seront donnés par les maires dans les villes, après l'avis des ingénieurs, et sous l'approbation du préfet, seront exécutés, jusqu'à ce que les plans d'alignemens aient été arrêtés au Conseil d'État, et, au plus tard, pendant deux années à compter de ce jour*(Duvergier, 1836, p. 297)

doute pas favorisé l'accélération des travaux de levé et de dessin des rues.

Le préfet Chabrol, revenu à la tête de la préfecture en 1815, commence à développer une politique d'aménagement urbain à l'échelle de la ville entière fondée sur des préoccupations esthétiques mais surtout fonctionnalistes : rendre les rues rectilignes pour réduire l'insalubrité, développer les trottoirs et le dallage basaltique, améliorer l'approvisionnement en eau par la multiplication des fontaines, améliorer la communication entre quartiers par des avenues et des places, etc. Ainsi, dans un rapport au Conseil Général du département de la Seine daté de 1819 et reproduit dans le tome II de ses *Recherches Statistiques sur la Ville de Paris et le département de la Seine* (De Chabrol de Volvic, 1823, p. 206), il met en avant l'importance des travaux d'alignement dans la capitale et leur accélération depuis 1817. S'il n'évoque pas les plans, nul doute que leur production subit le même changement de cadence. Toujours dans le même rapport (De Chabrol de Volvic, 1823, p. 195), il présente l'avancement des plans d'alignements : sur 1294 éléments de voirie environ²³, 1064 plans sont déjà levés, 220 restent à faire (les 10 plans manquants s'expliquent par le décompte approximatif du nombre des rues²⁴).

Les plans d'alignements constituent un outil de régulation puissant au service de la ville -tout propriétaire souhaitant construire un bâtiment doit consulter le plan de sa rue- mais un outil de gestion urbaine relativement mauvais. Très nombreux, chaque rue, place ou quai occupant une feuille, ils ne permettent pas une vision générale, pourtant essentielle pour une planification englobant tout Paris et ses faubourgs. C'était d'ailleurs déjà le constat de Verniquet lorsqu'il proposa au roi son projet de plan général. Un tel plan est le complémentaire essentiel des plans par rues : il donne une vision globale, sert d'index et permet de reporter sur lui tous les projets d'alignement ou de percement. Chabrol, sans doute très conscient de cet état de fait, met à profit le rattachement du bureau des plans de Paris à la préfecture de la Seine au début de l'année 1823 et la récupération de la seconde livraison du plan général de Verniquet jusqu'ici stocké au bureau du cadastre²⁵ pour lancer en 1824 un projet de grande envergure : la création d'un plan général d'alignements de la capitale en 48 parties, sur lequel devait être reportés tous les alignements existants ou en projet depuis Verniquet. C'est ce projet de grande ampleur qui sera à l'origine du dernier grand atlas de Paris réalisé à titre privé par un architecte : Théodore Jacobet

23. On compte alors 1070 rues, 120 cul-de-sacs, 34 quais et 70 places.(De Chabrol de Volvic, 1823, p. 195)

24. Chabrol dit lui même que *cette énumération diffère quelque peu de celle que l'on a faite pendant le cours de l'opération générale de dénombrement ; mais la question traitée dans [le] mémoire n'exige pas une précision absolue.*

25. Une note dans le tome VIII des *mémoires de la société de l'histoire de Paris et d'Île de France*, apporte quelques précisions (Société de l'histoire de Paris et de l'Île de France, 1881, pp. 301-303). D'après M^{me} de Chavagnac, fille de Verniquet, le plan est rassemblé en une feuille unique. D'après Verniquet lui-même dans ses mémoires, ce plan est en 6 parties), le plan est tout à fait abîmé par son usage intensif par le cadastre et les déplacements successifs qu'il subira jusqu'en 1817. Notons également que, comme le suppose (Pronteau, 1986), la première livraison du plan en 72 feuilles, dessinées et non gravées, était sans doute également en possession de la préfecture, ses géomètres -notamment Jacobet- déclarant régulièrement travailler avec "le grand plan de Paris en 72 feuilles". Ces deux livraisons ont aujourd'hui disparu.

1825 : Théodore Jacoubet et son premier plan de Paris

On voit apparaître Théodore Jacoubet au bureau des plans de la ville en 1823, au moment où celui-ci est intégré dans les bureau de l'Hôtel de Ville. Architecte de formation, il œuvre en qualité de dessinateur attaché à la confection des plans d'alignements. C'est dans ce cadre qu'il conçoit le projet de reprendre le grand plan de Verniquet qu'il juge désuet mais qui continue pourtant à être utilisé au sein de la préfecture. En 1823, Jacoubet s'associe avec M^{me} Gentil de Chavagnac, fille de Verniquet et qui détient la propriété intellectuelle de l'ouvrage. Ensemble, ils conçoivent d'abord le projet d'une nouvelle gravure avec mise à jour *mise au net et actualisée* du grand plan de Verniquet en six planches, réduites à l'échelle de 1/5.300^e. L'objectif est double : pour Jacoubet il s'agit de gagner sa place parmi les éditeurs de cartes de l'époque en tant que successeur de l'illustre Edmé Verniquet, et pour M^{me} Chavagnac de s'assurer la propriété d'une nouvelle version du plan de son père de haute qualité que la préfecture pourrait être amenée à utiliser pour remplacer l'ancien plan (étant à la préfecture, Jacoubet a accès à toutes les informations utiles). Réalisé avec l'aide de deux autres employés du bureau des plans de Paris, l'architecte Bailly et le commis dessinateur Mangot²⁶, le plan est prêt à être publié en 1825.

Cependant, comme nous l'avons évoqué dans la section 2.1 concernant l'atlas de Verniquet, la préfecture et Mme Gentil de Chavagnac sont à ce moment là en conflit, la propriété intellectuelle du plan revenant à cette dernière tandis qu'il est stocké à l'Hôtel de Ville qui en a l'usage. Puisqu'il est impossible de produire une nouvelle version de l'atlas de Verniquet sans l'aval de Mme Chavagnac, Jacoubet établit avec elle le 30 janvier 1825 un acte sous seing privé dans lequel il vend, pour 3000 francs, le plan gravé à Mme de Chavagnac qui se charge de le publier vers octobre 1825²⁷ sous le nom de *Plan général de la ville et des faubourgs de Paris : Année 1825*²⁸.

L'acte comporte une clause qui aura des effets indésirables sur l'atlas général. En effet, dans la 5^e clause Jacoubet s'engage à céder tous les droits de reproduction du plan et à ne pas *publier tout ou partie de ce plan, directement ou indirectement* (Gazette des Tribunaux, 1830, p. 1). Ainsi donc fin 1825 le plan mis à jour est publié et présenté au préfet Chabrol, mais Jacoubet a perdu tous droits dessus.

2.6.2 Création et première datation

Malgré l'interdiction qui le frappe, Théodore Jacoubet entâme, dès 1825, la création d'un atlas de Paris au 1/2.000^e qu'il veut être le successeur du grand atlas

26. Son nom est parfois orthographié Maingot, ou Manjon

27. On trouve l'annonce de publication dans *Bibliographie de France, volume 14, 1825* daté du 28 octobre 1825. (Journal Général de l'Imprimerie et de la Librairie, 1825, p. 728)

28. Le nom complet est : *Plan général de la ville et des faubourgs de Paris : Année 1825 / rapporté et dessiné par Mrs A. Mangot, attaché aux alignemens de Paris, Th. Jacoubet, architecte attaché aux mêmes alignemens, et Bailly, architecte. Basé d'après le grand plan trigonométrique de feu M. Verniquet, architecte conseiller du Roi; publié par madame sa fille; Dédié et présenté à Mr le comte de Chabrol de Volvic,...*; *Écrit par Martin; gravé par Adam* (Mangot et al., 1825). Notons que le plan a été présenté au préfet de la Seine à qui il est dédié, montrant que Jacoubet avait sans doute l'intention de le faire utiliser par les services de la préfecture

réalisé par Edmé Verniquet. Nous allons voir que cette réalisation n'est pas indépendante du plan introduit précédemment, mais aussi qu'elle va subir un grand nombre de contretemps qui feront de l'atlas général une source aussi incomparable que complexe d'utilisation.

Notes sur les sources documentaires

À partir d'ici, la majorité des informations dont nous disposons proviennent de témoignages de Jacobet lui-même ou de ses collègues dans divers procès les impliquant entre 1825 et 1842, relatés dans diverses revues judiciaires, et sur lesquels nous reviendrons au fur et à mesure du propos. En particulier, il s'agit de la *Gazette des Tribunaux*²⁹, de *Journal des débats politiques et littéraires*³⁰ et de la *Gazette judiciaire*³¹.

Il s'agit de journaux quotidiens dans lesquels les débats tenus au cours d'audiences judiciaires à la cour d'assises de la Seine sont retranscrits par des journalistes y assistant. Il ne s'agit donc pas de sources d'informations aussi fiables que l'on pourrait le souhaiter, parfois approximatives et souvent tronquées. Les trois gazettes précitées ont cependant l'avantage de rapporter les mêmes faits et les mêmes propos, tout en étant alimentées par des journalistes différents. En conséquence, si l'on peut difficilement donner foi à une seule gazette, le croisement de ces trois sources permet de réduire les doutes sur les informations relatées. Nous jugeons alors correcte une information dès lors qu'elle est présente dans au moins deux des trois journaux.

1825-1836 : Publication de l'atlas général

Dans le témoignage de Jacobet à la cour d'assises de la Seine le 11 novembre 1842 (voir 2.6.4, il apparaît que *le manuscrit*³² [de l'atlas] *était terminé en 1825* et qu'un premier prospectus a paru *en 1826* (Gazette des Tribunaux, 1842a, p. 2). Le prospectus³³ annonce la parution de l'atlas en ces termes :

Atlas général de Paris, de ses faubourgs et monumens. Par Th. Jacobet, architecte. Cet atlas aura 54 feuilles, qui paraîtrons en 9 livraisons de 6 feuilles, de trois en trois mois.. Le contenu annoncé est le suivant :

- *tous les monuments publics détaillés avec le plus grand soin ;*
- *les maisons particulières avec leurs jambes-étrières et leurs numéros ;*
- *tous les projets de percements, ainsi que les alignements arrêtés par ordonnance royale ;*
- *la division intérieure des propriétés qui sont atteintes par les projets définitivement arrêtés ou qui les avoisinent immédiatement ;*
- *de plus, les routes royales, les ruelles et les chemins vicinaux seront tracés dans toute l'étendue que contiendra le cadre qui environne le plan.*

Comme pour le plan des 48 quartiers de Vasserot et de nombreux autres plans de l'époque, il s'agit donc d'une vente par souscription, le vendeur faisant la publicité

29. <http://data.decalog.net/enap1/liens/Gazette>

30. <http://gallica.bnf.fr/ark:/12148/cb39294634r/date>

31. <http://gallica.bnf.fr/ark:/12148/cb327812934/date.r=Gazette+Judiciaire.langEN>

32. c'est à dire le plan minute

33. Ce prospectus existe toujours, voir Jacobet (1826)

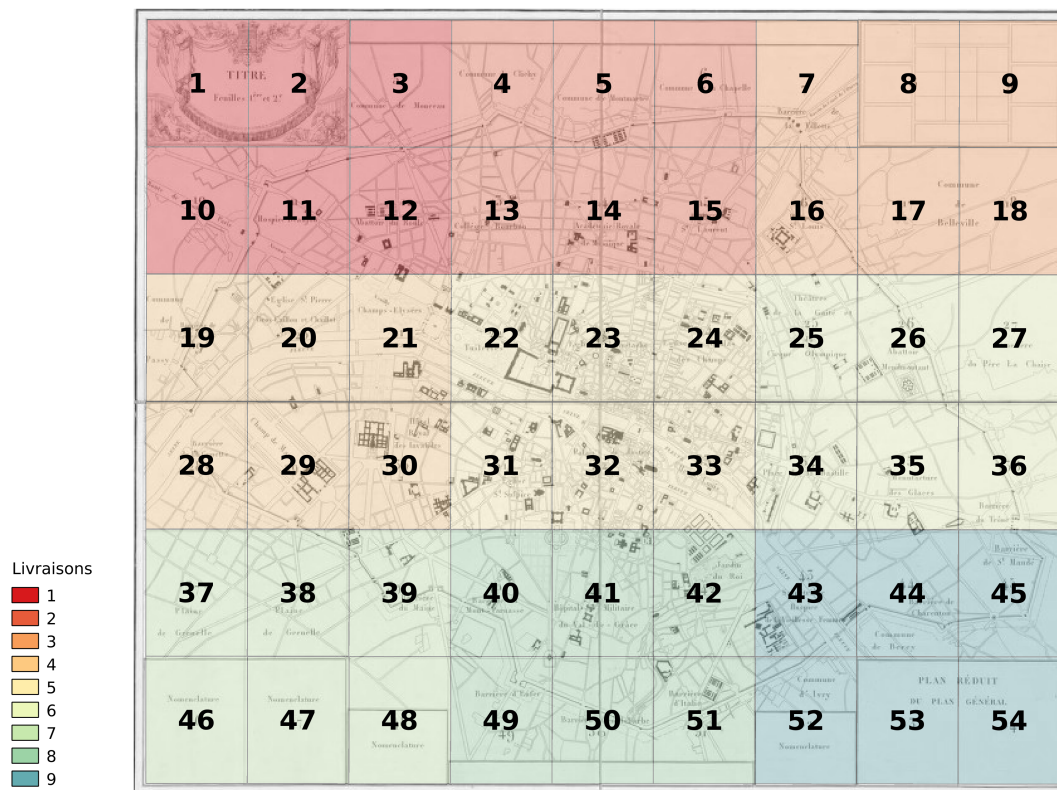


FIGURE 2.2 – Index des planches de l'atlas général de Jacoubet, tracé sur le plan d'assemblage, qui en constitue les feuilles 8 et 9.

du plan avant sa gravure, puis livre par lots les parties du plan aux souscripteurs. Si la souplesse offerte par la souscription permet à l'auteur de vendre un plan en cours et de financer les nouvelles planches avec l'argent récolté, elle complique le suivi de la construction du plan, et met en danger sa cohérence : un trop grand écart entre la réalisation de deux planches adjacentes va obliger à reprendre et mettre à jour la première. Ainsi, le manuscrit semblait déjà terminé en 1825, et le processus de gravure et de publication fut enclenché l'année suivante. Le plan d'assemblage de l'atlas, illustré dans la figure 2.2, donne l'index des différentes planches qui le constituent lorsqu'il paraît dans sa version finale.

La gravure réalisée par Antony-Grégoire Niquet est efficace puisque le 13 janvier 1827 paraît la première livraison du nouvel atlas³⁴, suivie dans le courant de l'année par deux autres livraisons. Bien que nous n'ayons trouvé trace des autres livraisons, rien ne laisse croire que le rythme ait baissé jusqu'en fin d'année 1828. Avec un rythme de trois livraisons par an, à la fin de l'année 1828 six livraisons du plan auraient été gravées et publiées.

Pourtant, cette publication n'est pas sans incident. En effet, dès 1826 et l'apparition du prospectus, M^{me} Chavagnac, soupçonnant Jacoubet de publier une nouvelle

34. En réalité, il s'agit de la seconde livraison prévue par Jacoubet, soient les feuilles 4, 5, 6, 13, 14 et 15, la première ayant pris du retard. Cette livraison est la seule qui soit parvenue jusqu'à nous dans son état originel. Elle se trouve actuellement à la Bibliothèque Historique de la Ville de Paris. (Jacoubet, 1827)

version du plan de Verniquet qu'il lui avait vendue l'année précédente, l'assigne devant le tribunal de première instance de la Seine pour non respect du contrat sur le plan de 1825. En effet, il n'échappe certainement pas à M^{me} Chavagnac qu'il est difficilement envisageable qu'un nouveau plan de Paris, à une échelle aussi grande, ait pu être entièrement réalisé en quelques mois là où 6 ans avaient été nécessaire à Verniquet et 14 ans à Vasserot, d'autant plus qu'il publie ce nouveau plan à la suite de celui qu'il lui a cédé. De plus, l'un des auteurs du plan, Mangot, avait entre temps vendu frauduleusement les 72 feuilles dessinées de l'atlas de Verniquet en possession du bureau des plans³⁵, jetant ainsi quelques doutes sur les pratiques des géomètres et architectes du bureau.

Le 20 décembre 1828 M^{me} Chavagnac obtient gain de cause et Jacobet se voit condamné à la cessation de la publication et à la destruction des minutes et des cuivres. Faisant appel, une commission d'enquête est nommée le 7 août 1829 pour étudier le plan et décider s'il s'agit d'une œuvre originale ou d'une nouvelle publication du plan de 1825, mais, en attendant son rapport, l'interdiction de publication est maintenue. Finalement, le 8 mars 1830 la commission conclut que :

- L'atlas n'est pas une version à une autre échelle du plan de Verniquet.
- Ce n'est pas non plus une nouvelle version du plan réduit publié en 1825.
- L'atlas a été effectué sur une nouvelle triangulation et, par conséquent, il est totalement original.

Jacobet sort donc vainqueur de ce procès et reprend aussitôt la publication de son atlas. Il ne rencontrera plus de problèmes majeurs, hormis une soustraction momentanée des cuivres par son graveur Niquet en 1833, ce qui l'incitera à changer de graveur et à choisir Louis-Marin Bonnet pour le remplacer (*Gazette des Tribunaux*, 1833, p. 4). La gravure de toutes les planches semble avoir été terminée autour de 1834, Jacobet affirmant lui-même qu'à cette date son atlas *était terminé, ou presque* (*Gazette des Tribunaux*, 1842a, p. 2). Finalement, il publiera courant 1836 la totalité de son plan sous la forme d'un grand livre, auquel il ajoutera les éléments suivants :

- Les plans de Paris à diverses époques, très inexacts : *sous la domination Romaine, sous le règne de François I^{er}, sous le règne de Louis XIII, sous le règne de Philippe Auguste et sous Louis Philippe I^{er}.*
- Les plans des cimetières de Paris (incorporés au plan).
- Le plan résumant la triangulation ayant permis le levé de l'atlas.

1836 : publication de l'atlas et critique de la datation

Ainsi donc, l'atlas de Jacobet semble, tout du moins dans sa phase de levé et de gravure, plus ancien que ce qui est généralement avancé (1827-1840 pour Jeanne Pronteau (Pronteau, 1986), 1827-1839 pour Pierre Pinon (Pinon *et al.*, 2004, p. 94)). Le relevé du plan daterait donc de la période 1824-1826, et sa gravure s'échelonnerait principalement entre 1827 et 1834. Finalement, le plan sera publié sous forme d'atlas -c'est à dire d'un livre relié contenant toutes les planches- en

35. Fort heureusement, la préfecture récupérera les plans avant qu'ils soient vendus et Mangot sera limogé en 1831.

1836, tel que l'indiquent effectivement les planches 1 et 2³⁶ et comme en témoigne également Jacoubet (*Gazette des Tribunaux*, 1842a, p. 2).

Cependant, la datation que nous avons établie amène une seconde question : si le plan a effectivement été tracé vers 1825, le rythme des transformations de Paris le rendent déjà obsolète à sa publication en 1836. Il paraît peu vraisemblable que Jacoubet, qui voulait faire de son atlas le nouveau *Grand Plan de Paris* dans lequel tous les projets de voirie figureraient, ait publié le plan daté de 10 ans sans l'avoir jamais mis à jour, y avoir reporté les nouveaux projets et percements. Il publiera d'ailleurs plusieurs fois son plan dans les années suivantes en grand format (1838,1841) ou en réduction au 1/10.000^e dès 1836. Il en fera une mise à jour complète en 1841. On trouve une version gravée des trois premières livraisons en 1833, soit 6 ans après leur première parution (Jacoubet, 1833)³⁷. Si l'atlas a été modifié en cours même de publication, la chronologie établie jusqu'ici est alors tout à fait insuffisante pour connaître le véritable temps de réalisation du plan. Plus encore, il nous faut déterminer si le plan a subi des mises à jours globales, où si chaque planche a suivi une trajectoire différente. Seule une seconde analyse fondée sur le contenu de l'atlas peut répondre à ces questions.

2.6.3 Datation fine

Datation par planches

Afin de déterminer l'intervalle pendant lequel chaque planche a été dessinée puis gravée et ainsi vérifier si les livraisons sont cohérentes ou s'il y a eu des mises à jour ponctuelles entre 1827 et 1836, nous avons analysé le contenu de chacune des planches de la première édition de l'atlas datée de 1836. La méthodologie adoptée pour fixer les bornes de chaque intervalle est la suivante :

- La borne inférieure peut être approchée en connaissant les dates de création des éléments représentés. Par exemple, si un alignement est tracé sur une planche alors elle a été gravée après la date de l'ordonnance d'alignement.
- La borne supérieure est trouvée de la façon inverse. Ainsi, lorsqu'une décision de voirie est prise et non représentée sur la planche, cela signifie que cette dernière est antérieure. Cette étape est cependant plus fragile car elle est sensible aux éventuels oublis de l'auteur, mais aussi car certains percements sont omis car réalisés sans autorisation officielle (C'est le cas en particulier à l'est de Paris avec les rues Gambrey et Pierre Levée, mais aussi parfois à l'ouest comme pour la rue de Milan).

Il est aussi nécessaire de choisir quelles informations sont suffisamment fiables pour guider la fixation des bornes temporelles. Ici, nous nous sommes appuyé sur le prospectus de Jacoubet qui donnait une importance particulière à la représentation des ordonnances de voirie. Ainsi, à partir des objectifs du plan d'origine, nous avons choisi de nous appuyer sur les éléments présentés dans le tableau ci-dessous.

36. Jeanne Pronteau note cependant que cette date est sans doute fictive, les armoiries dessinées sur la même planche étant celles de la Restauration.

37. Cette version, conservée à la BNF (Jacoubet, 1833), ne contient cependant que les planches 3 à 7, 10 à 12 et 16 à 18. Il manque donc les planches 1,2, 8, 9 ,13, 14 et 15.

Alignements et percements	Rappelons que le plan rapporte <i>tous les projets de percements, ainsi que les alignements arrêtés par ordonnance royale</i> . Ainsi, tous les élargissements, projets de percements et d'ouvertures sont représentés sur le plan dès lors qu'ils sont décidés par ordonnance royale ou décret, dont les dates exactes peuvent être retrouvées.
Dénominations	Les changements de noms des rues et bâtiments, lorsqu'ils résultent de décisions royales ou préfectorales, sont reportés sur le plan. Cette indication est précieuse pour un atlas qui a traversé une révolution ayant amené divers changements de dénominations des rues (Lafayette, place du Panthéon, etc.)
Bâtiments publics	Les bâtiments publics représentés sur le plan sont soit existants, soit parties de projets officiels dont la date de décision peut également être retrouvée.

TABLE 2.1 – Types d'événements affectant Paris ayant permis de dater les planches de l'atlas de Jacobet, triés par ordre décroissant de fiabilité.

Notons que, puisque le plan rapporte les projets officiels et non ce qui est déjà effectué, il peut nous induire en erreur. Par exemple, les plans de l'Hôtel de Ville et de la Bibliothèque de Roi sont tracés bien qu'ils n'aient jamais été réalisés en l'état. De même, certaines rues telle celle de Munich, partie du lotissement de l'Europe, sont dessinées mais n'ont jamais été ouvertes.

Nous avons choisi, pour dater les rues, projets et ordonnances, de nous appuyer uniquement sur des sources officielles :

- Les *Dictionnaires Historiques des Rues de Paris* par les frères Lazare, datés de 1844 et 1855 (Lazare et Lazare, 1844, 1855).
- Le *Recueil général des lois et des arrêts* contenant les ordonnances royales
- La nomenclature des rues de Paris diffusée en 2012 par la Mairie de Paris³⁸.

Nous avons en outre décidé de ne pas nous appuyer sur d'autres plans contemporains qui, par comparaison, auraient permis de préciser les dates des rues. En effet, la plupart des autres plans de l'époque sont très peu fiables, parce que, face au rythme de transformation de Paris et pour éviter une trop rapide obsolescence, ils représentent en avance de nombreux percements. Il est courant qu'un plan contienne le tracé de rues ayant été créées plusieurs années après sa parution, ou bien même de rues jamais percées. C'est le cas des plans de Girard (Girard, 1830), Achin (Achin *et al.*, 1825) ou Tardieu (Tardieu, 1838). Seule la seconde édition de l'atlas de Jacobet, publiée vers 1839, a été utilisée pour discriminer les projets de voirie décidés avant 1836 mais pas encore effectués en 1839.

Le résultat de cette analyse est présenté au sein du tableau 2.2, qui fait apparaître pour chaque planche l'intervalle temporel de son dessin, que nous obtenons à partir d'une somme d'indices temporels rassemblés dans l'annexe G.

Pour chaque planche, l'intervalle temporel  décrivant la période de son possible dessin est déduit d'une somme d'événements datés concernant l'urbanisme

38. <http://parisdata.opendatasoft.com/explore/dataset/voiesactuellesparis2012>

parisien : percement de rues, alignements, construction, etc. Les types d'événements sont décrits dans le tableau 2.1, et la liste complète est fournie dans l'annexe G. Ces événements sont de deux types. Les premiers notés ► sont dessinés sur la planche et sont donc considérés comme antérieurs au dessin. Les seconds, notés ◄, ne sont pas représentés sur la planche et sont donc considérés postérieurs au dessin. Ces deux catégories d'événements permettent de définir les bornes inférieures et supérieures de l'intervalle temporel pendant lequel la planche a pu être dessinée puis gravée. L'écart temporel entre un événement, généralement daté au jour près, et son report sur la carte nous est inconnu, produisant une incertitude sur la position des bornes au sein d'une année. Cette incertitude est représentée dans le tableau par la couleur ■. Enfin, certains événements se contredisent, ce qui peut arriver dans le cas de corrections partielles d'une planche ou d'oublis. Ils se trouvent indiqués par la couleur ■. Les bornes de certaines planches n'ayant pas pu être déterminées par manque d'informations, en particulier dans les zones périphériques, nous avons arbitrairement fixé un minimum et un maximum, respectivement 1825 et 1836, la première date correspondant au début de la réalisation du grand atlas et la seconde à son année de publication. Seules les planches ayant un contenu cartographique sont représentées.

TABLE 2.2 – Datation des planches de l'atlas de Jacoubet commencé en 1825 et publié en 1836, rassemblées par livraison.

Planches	Années											
	1825	26	27	28	29	30	31	32	33	34	35	36
3	■	■	■	◄	■	◄	■	■	■	■	■	■
10	►	■	■	■	■	■	■	■	■	■	■	■
11	►	■	■	■	►	■	■	■	■	■	■	◄
12	■	►	■	■	■	■	■	■	■	■	■	◄
4	■	►	■	■	■	■	■	■	■	■	■	◄
5	►	►	■	■	■	■	■	■	◄	■	■	■
6	■	►	►	■	■	■	■	■	■	■	■	■
13	■	►	►	■	■	◄	◄	■	■	■	■	■
14	■	►	■	■	■	►	►	◄	◄	■	■	■
15	■	■	►	■	►	■	■	◄	■	■	■	◄
7	■	■	■	■	■	►	■	■	■	■	■	■
16	►	►	►	■	■	►	■	■	■	■	■	◄
17	■	■	■	■	■	■	■	■	■	■	■	■
18	■	■	■	■	■	■	■	■	■	■	■	■
19	■	■	■	■	■	■	■	►	■	■	■	■
20	■	■	■	■	■	■	■	►	►	■	■	■
21	■	■	◄	►	►	►	►	■	■	■	■	◄
28	■	►	■	■	■	◄	■	■	■	■	■	■
29	►	■	■	■	■	■	■	■	■	■	■	■
30	■	■	►	■	■	■	■	■	■	■	■	■
22	■	■	■	►	■	■	■	►	■	◄	■	■
23	■	■	■	■	■	■	■	■	►	■	■	◄
24	■	◄	◄	►	►	◄	◄	■	■	■	■	◄

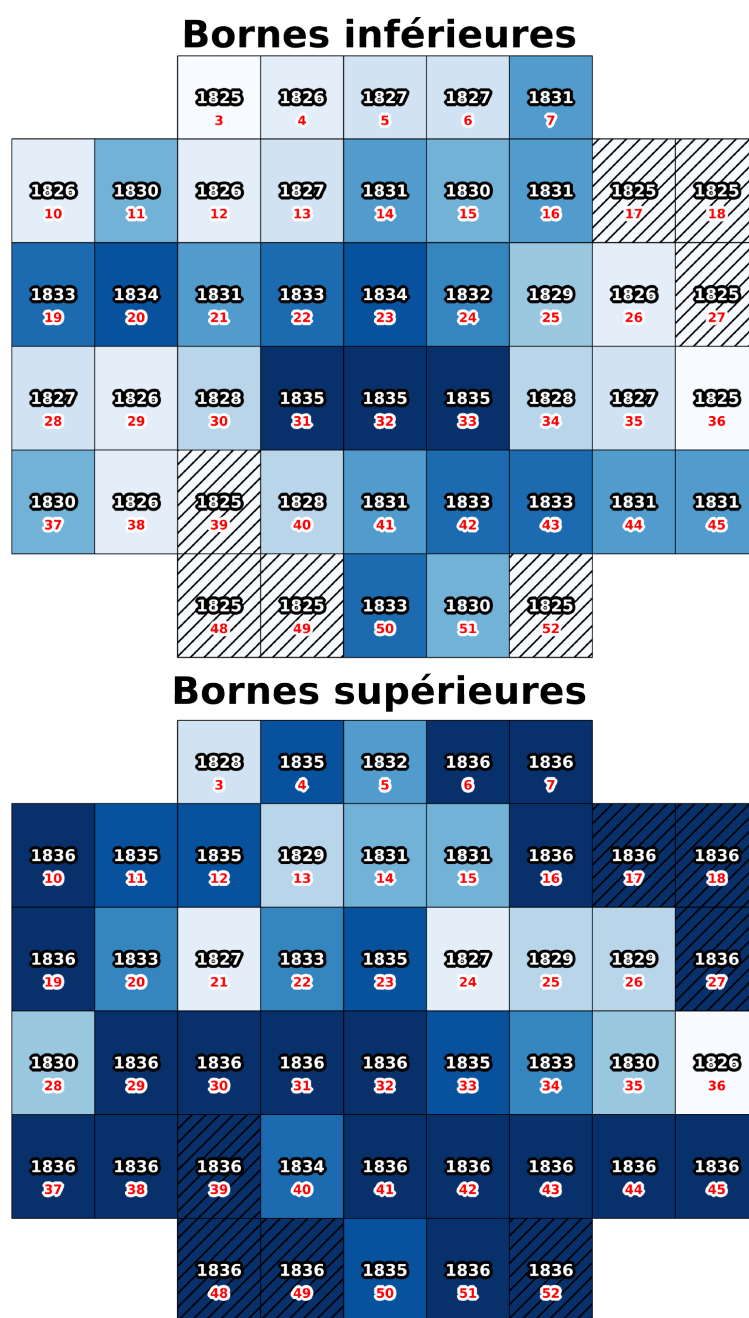
La situation n'est donc véritablement chaotique que pour les quatre premières livraisons et la sixième. Les autres livraisons suivent une trajectoire certes décalée de plusieurs années par rapport au programme de Jacoubet, mais qui reste cohérente. Notons cependant qu'entre 1827 et l'édition de 1836, 9 ans se sont écoulés pendant lesquels l'architecte publiait les livraisons indépendamment les unes des autres et non rassemblées dans un atlas. Neuf années durant lesquelles les opérations de voirie se sont poursuivies. Ces livraisons, destinées surtout à financer l'atlas en cours de préparation, étaient donc constituées de planches appelant rapidement des mises à jour, en particulier pour être rassemblées dans un ouvrage synthétique. Nous posons donc l'hypothèse de mises à jour successives plutôt que la publication d'un atlas issu d'un levé unique. Au cycle unique levé-dessin-gravure se substituerait donc un enchaînement de ces cycles à l'échelle d'une livraison ou de quelques planches. De plus, cette hypothèse permet de donner une explication aux diverses planches présentant une incohérence temporelle. En effet, cette incohérence peut être le fruit d'une mise à jour partielle d'une planche, par exemple avec l'ajout ponctuel d'un projet de voirie.

Deux éléments nous permettent d'appuyer cette hypothèse. Tout d'abord, on a constaté que certaines planches des quatre premières livraisons étaient postérieures aux autres planches de la même livraison. Il faut ici se rappeler que ces livraisons (sauf la quatrième) datent du début de la publication en 1827, et se situent avant le procès avec M^{me} Chavagnac, avant la révolution de 1830 et avant la réorganisation du bureau des plans en 1831. Elles étaient donc largement obsolètes à la publication des dernières livraisons, celles-ci étant alors plus proches du Paris de 1835. Il n'est donc pas absurde de penser que Jacoubet ait repris ces planches et les aient corrigées. La figure 2.3 fait apparaître un second phénomène significatif. On constate dans cette figure que les dates des bornes inférieures sont corrélées, en plus de la livraison, à la zone de Paris concernée. En effet, les planches du centre sont régulièrement plus tardives que les livraisons dont elles font partie : Jacoubet aurait donc mis à jour les planches de son atlas dans les zones centrales, plus fortement soumises aux transformations (en particulier, aux alignements). Notons cependant que ces planches sont également les plus riches en indices et par conséquent les plus simples à dater, les périphéries étant alors essentiellement marquées par les grands projets de lotissement commencés en 1820-1825.

Conséquences

La conséquence première de ces constatations est la remise en question totale de l'atlas de Jacoubet comme un objet homogène qui pourrait être traité comme une représentation globale et fiable de Paris en 1836. Tout autant que les précédents plans, celui-ci n'est pas une "tranche temporelle" de Paris mais un mille-feuille qui accumule des fragments locaux d'une ville qui change sans cesse. Plus encore, l'hétérogénéité des mises à jours entraîne une décohérence des différentes parties du plan. Ainsi, plusieurs feuilles de l'atlas pourtant adjacentes ne sont pas cohérentes entre elles. Un des exemples les plus parlants est illustré dans la figure 2.4 : la place du Marais, qui contiendra à partir de 1833 l'entrepôt réel des douanes³⁹, présente sur

39. Ordonnance royale du 28 juin 1833. Voir Guével (2004)



Légende

- Bornes temporelles inférieures et supérieures 1830 Date de la borne
- Bornes temporelles inférieures et supérieures 50 Numéro de la planche

FIGURE 2.3 – Bornes temporelles inférieures et supérieures reportées sur l'emprise de chaque planche de l'atlas général de Jacobet.

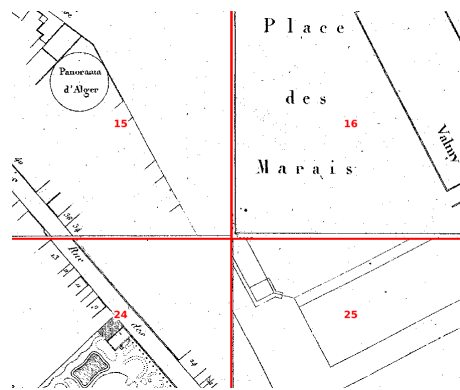


FIGURE 2.4 – Un exemple d’incohérence entre quatre planches de l’atlas de Jacoubet, due à des périodes de dessin différentes.

4 planches de l’atlas, possède à chaque fois une représentation différente. Datée de 1821 et entourée de trois rues, elle est dans son état d’origine sur les planches 15 et 16, mais sur la planche 24 elle est absente. Finalement, seule la planche 25 contient les tracés de parties de l’entrepôt, construit à partir de 1833. De telles incohérences se retrouvent régulièrement, en particulier dans les zones en pleine mutations au moment de la réalisation du plan (en particulier les quartiers Saint-Lazare, Europe, Poissonnière, Grenelle, canal Saint-Martin et Austerlitz).

2.6.4 Jacoubet et l’administration parisienne : l’insertion de l’atlas général dans la gestion de la Ville

Dans la section précédente, nous sommes parvenus à reconstituer la genèse de l’atlas de Jacoubet, montrant ainsi qu’il s’agit d’un atlas particulier, non seulement en raison de son ambition et de ses dimensions, mais aussi par les liens que l’on voit se tisser entre l’atlas, son auteur et l’administration de la Seine. Contrairement à la plupart des plans de l’époque (la première moitié du XIX^e siècle voit en effet proliférer les plans de Paris, plus ou moins exacts), celui de Jacoubet a été réalisé par un architecte qui travaillait à l’endroit même où la voirie parisienne était administrée. Étant donnée la précision et l’échelle du plan qui en font un concurrent sérieux du cadastre de Vasserot et de l’atlas de Verniquet, on peut se demander si ce n’est pas ce plan qui a servi à la préfecture jusqu’à l’arrivée de Haussmann, faisant de lui le ”plan officiel” qui précède celui de Haussmann. Dans ce cas, le plan de Jacoubet constituerait un des ”chaînon manquant” entre les plans de Vasserot et celui de Haussmann. Ainsi, nous manipulerions -à part pour Maire- uniquement des plans qui ont servi à modeler la ville de Paris, témoins privilégiés des projets et visions des administrateurs de la Seine.

Dans cette section, nous allons essayer de montrer que cet atlas n’est pas seulement une œuvre singulière, mais surtout qu’il est un révélateur du fonctionnement de la préfecture de la Seine de 1820 à 1842 et qu’il ouvre une fenêtre sur une ville où la frontière quasi-inexistante entre les fonctionnaires et les entrepreneurs privés profite largement aux intérêts des uns et des autres. Nous allons également pouvoir donner

quelques pistes permettant de comprendre comment il est possible que Jacobet ait construit un atlas au 1/2.000^e, sur une nouvelle triangulation, alors même qu'il travaillait à la préfecture.

Afin d'explorer cette hypothèse, nous allons nous appuyer sur une affaire judiciaire datant de 1842, appelée par ses contemporains *Affaire Hourdequin*, dans laquelle une partie des employés de la préfecture chargés de la réalisation des plans d'alignements se sont vu accusés de corruption et de vol de plans. Cette affaire est devenue un véritable scandale lorsqu'elle a mis au jour la profonde corruption qui frappait les bureaux chargés des alignements, précipitant rapidement la plupart des responsables de ces bureaux, dont Théodore Jacobet, dans un procès qui durera 2 ans (18 novembre 1840-20 novembre 1842), demandera le témoignage de 132 personnes et aura d'importantes répercussions sur le fonctionnement de la préfecture et sur la réorganisation des services adoptée plus tard par Haussmann.

Notre objectif n'est pas de dérouler entièrement le récit des 13 jours de procès, du 7 novembre au 20 novembre 1842, qui constituent le cœur du scandale et pendant lesquels seront faites la quasi-totalité des révélations sur les bureaux de la préfecture. Il s'agit plutôt d'en extraire les éléments qui nous permettront de reconstruire le contexte dans lequel le plan de Jacobet a vu le jour, et d'entrevoir les raisons qui font de lui l'ancêtre des plans officiels de Paris.

Le bureau des plans de Paris

Afin de bien comprendre les relations entre les différents personnages de l'affaire, mais aussi le statut de Théodore Jacobet au sein de la préfecture de la Seine, il faut savoir comment étaient organisés les services gérant la préparation et la réalisation des plans d'alignements et de percements, ainsi que ceux qui avaient pour charge l'application des amendes pour non respect des réglementations de voirie. Comme nous l'avons dit, l'affaire Hourdequin va causer une totale réorganisation des bureaux de la préfecture. Nous nous intéressons donc ici uniquement à la période qui précède, de 1823 - date d'arrivée de Jacobet- à 1842.

La figure 2.5 donne les différentes organisations de la Grande Voirie de Paris, depuis le rattachement du bureau des plans jusqu'à la réorganisation de Haussmann. Un *bureau* est une unité administrative qui correspond aujourd'hui à un *service* ; composé d'un chef, d'un sous-chef et d'un nombre variable d'employés, il se voit chargé d'une tâche précise. Ainsi, le *bureau des plans* a en charge la création des plans d'alignements de Paris, leur communication au public, au préfet et au conseil municipal, ainsi que la création de tout plan nécessaire à l'établissement d'amendes, d'expropriations, etc. C'est enfin ce bureau qui se voit chargé par Chabrol de la confection du grand plan de Paris en 48 parties. Notons également que jusqu'en 1831, le bureau des plans était une annexe de la Grande Voirie et non directement une de ses composante, ce statut lui conférant une relative autonomie.

C'est dans ce bureau que Jacobet entre en 1823 et bâtira toute sa carrière sous les ordres de Chabrol puis de Rambuteau, son statut évoluant en fonction des événe-

ments. Ainsi, entré comme simple architecte en 1823, il deviendra sous-chef adjoint en 1827, puis sous-chef en 1833. Il finira par devenir chef du bureau en 1839 avant de redevenir sous-chef en 1842. Il est de plus *vérificateur* depuis 1827, ce qui signifie qu'il contrôle, vérifie les plans réalisés par les employés. Sans son visa et sa signature, le travail est rejeté. Jacoubet n'est donc pas un simple employé, mais il joue un rôle important dans ce service alors qu'au même moment paraît son atlas. Nous avons dit que l'affaire Hourdequin permettait de replacer l'atlas de Jacoubet dans le contexte de la préfecture de la Seine, mais aussi plus largement dans celui des pratiques des ingénieurs cartographes de Paris durant la période précédant Haussmann. Pour reconstituer la matrice qui a permis la création de cet atlas, nous résumons le *procès Hourdequin*.

Les causes de l'affaire

Le scandale du bureau des plans éclate le 18 novembre 1840 lorsque l'architecte Adolphe Morin, attaché au Cercle du Jeu de Paume du passage Sandrié, est arrêté pour avoir dérobé un billet de 1000 francs à un certain Nanteuil. Arrêté, l'enquête menée sur ses antécédents fait rapidement surgir plusieurs faits étranges : ancien employé du bureau des plans à la préfecture de la Seine, il a été limogé en 1834 pour soustraction de plans⁴⁰ et imitation des signatures de Auguste-Jean Hourdequin et Théodore Jacoubet, respectivement chef de la Grande Voirie et sous-chef du bureau des plans. Sur ces révélations, la cour d'assises de la Seine nommera une commission dirigée par Héricart de Thury chargée d'enquêter sur le bureau du plan. Pendant un an les experts de la commission vérifieront plus de 8000 pièces (plans, lettres, etc.), découvrant alors que de nombreux plans sont manquants, et que certains employés sont manifestement corrompus. En conséquence, le juge décide le 2 juillet 1842 de la mise en examen de 5 employés de la Grande Voirie pour les motifs suivants :

- Alexandre Philidor (employé) : Accusé d'avoir soustrait des plans servant à mettre en place une contravention, rendant celle-ci inapplicable, et d'avoir reçu un "pot-de-vin" de la part du contrevenant.
- Nicolas Boutet (employé) : Accusé d'avoir détruit des contraventions, ou délibérément ralenti leur application en échange de dons en nature ou en argent.
- Auguste-Jean Hourdequin : Accusé d'avoir abusé de sa position pour spéculer à son avantage sur des terrains et des propriétés menacés de démolition pour percement ou reculement, d'avoir favorisé certains entrepreneurs lors d'appels d'offres publics et enfin d'avoir accepté des dons d'argent en échange de la disparition de contraventions.
- Étienne-Alexandre Solet : Accusé de la soustraction de 5 plans, dont celui, officiel, d'une rue *projetée du Louvre à la Bastille*⁴¹
- Adolphe Morin : Accusé de faux en 1832 et 1834 et de la soustraction des plans-minutes⁴² de la rue de la Verrerie et d'un terrain *compris entre le Quai d'Orsay, les avenues de Suffren et de Lowendal, et le boulevard extérieur.*,

40. c'est à dire vol ou recel de feuilles ou minutes des plans produits à la préfecture.

41. Il s'agit sans nul doute du projet de la rue Impériale, prévue dès le premier Empire, qui devait relier la place du Louvre à celle de la Bastille. Jamais réalisé, le projet sera finalement abandonné au profit du percement de la rue de Rivoli.

42. Le plan-minute d'un alignement ou de toute autre opération de voirie est un document sensible, puisqu'il est fait très en amont de son exécution et de sa publicité par ordonnance ou décret.

1823 : Jacoubet arrive à la préfecture

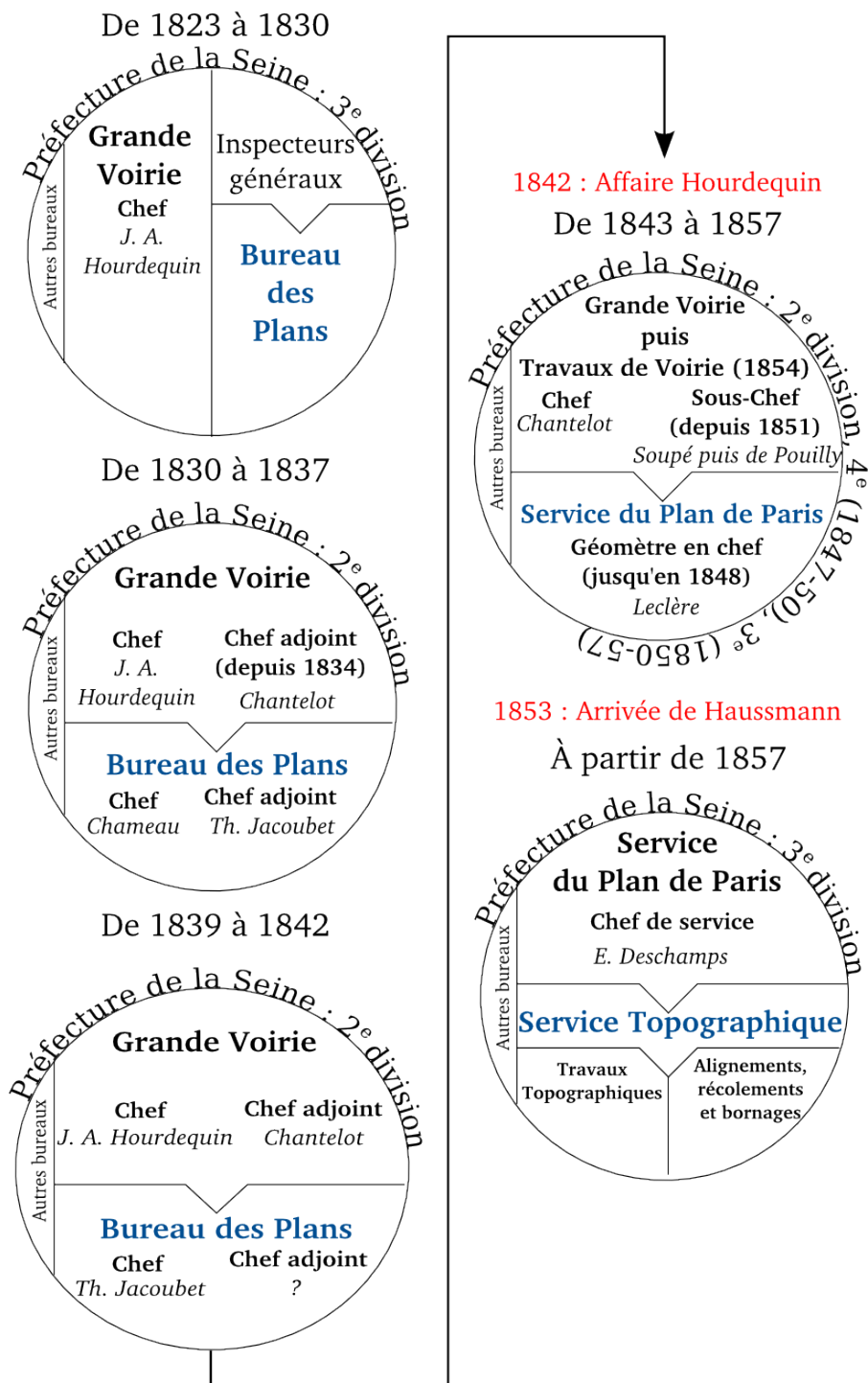


FIGURE 2.5 – Organisation du Bureau des Plans de Paris, depuis son déplacement du ministère de l'intérieur à la préfecture de la Seine jusqu'à sa réorganisation complète par Haussmann. Chaque ligne rouge indique une phase de réorganisation majeure, ayant des conséquences sur la production et l'utilisation des plans de Paris à la préfecture. Les couleurs indiquent la filiation du bureau des plans.

ainsi que du plan officiel de la rue des Lombards.

Jacoubet n'est pas accusé, bien qu'il soit placé momentanément en prévention *pour négligences*. Il comparaitra cependant régulièrement en qualité de témoin et d'expert cartographe, puis il prendra activement part aux débats et répondra dans une lettre ouverte aux attaques de l'avocat Gustave Louis Chaix d'Est-Ange (Jacoubet, 1842). Plutôt que de nous attarder sur chaque personnage, nous allons passer en revue les éléments d'accusation qui forment le canevas de l'entreprise de Jacoubet. Ces éléments, au nombre de 3, ne font pas nécessairement partie des faits incriminés lors du procès, soit parce qu'ils sont anciens, le tribunal décidant alors de les ignorer, soit parce qu'il s'agit d'accusations morales que la justice ne saurait traiter. Il s'agit des éléments suivants, qui feront l'objet de points détaillés car ils permettent de mieux comprendre le rôle joué par Jacoubet :

- Les travaux personnels effectués par les agents du bureau sur leur temps de travail, et l'utilisation de prête-noms pour le financement de ceux-ci.
- Le vol de plans, la redondance de travaux (parfois faits par la même personne).
- Les actions spéculatives menées en association avec des entrepreneurs privés.

Les travaux privés au sein du bureau des plans de Paris.

Jusqu'en 1831, il existait trois groupes d'employés. Les premiers, dont Jacoubet fait partie, étaient des salariés à plein temps percevant un appointement. Les seconds étaient attachés temporairement au bureau et recevaient également un salaire fixe. Enfin, les derniers étaient payés à la tâche. Bien que les postes fixes offrent une plus grande sécurité, ils présentent l'inconvénient de gêner les travaux privés d'employés qui exercent également en parallèle un métier d'architecte ou de géomètre puisqu'il leur est interdit d'y travailler pendant les heures dues à la préfecture. Ainsi donc, la pratique courante (Gazette des Tribunaux, 1842b, p.2-3) consistait à s'entendre avec les employés payés à la tâche, qui servaient de prête-noms. Tolérée pendant un temps, cette pratique sera massivement utilisée par Jacoubet pour réaliser son atlas. En effet, il apparaît dans le cours du procès que l'accusé Solet, qui jusqu'en 1832 sera dessinateur pour le bureau des plans, avait prêté son nom pour certaines tâches réalisées en fait par Jacoubet. En particulier, il a travaillé comme dessinateur sur l'atlas au moins jusqu'en 1829 et dessinera, d'après Jacoubet, près des deux tiers du plan (Gazette des Tribunaux, 1842a) qu'il effectuera dans son atelier privé et non à l'Hôtel de Ville.

Jacoubet, dans son témoignage le 11 novembre 1842 (Gazette des Tribunaux, 1842a), annonce ainsi qu'il a fait travailler sur son plan deux personnes extérieures à la préfecture, dont on ne connaît pas les noms, recrutées pour l'occasion, ainsi que deux employés : Solet et un ingénieur civil nommé Cannot-Tissier. Ces pratiques, si elles semblent tolérées par Chabrol, vont disparaître en 1831 sous la direction du préfet de Bondy. Celui-ci, pour faire stopper ces agissements qui nuisent grandement à la réalisation du grand plan voulu par Chabrol, nomme en 1831 Hourdequin à la tête de la Grande Voirie, lequel s'empresse d'interdire le démarrage de nouveaux travaux privés, mais il autorise la poursuite de ceux en cours. Un témoignage de l'accusé Morin indique d'ailleurs que Jacoubet a continué à utiliser cette pratique après 1830. Si la fiabilité des propos de Morin, renvoyé de la préfecture par Jacoubet en 1834, est presque nulle, il est tout de même possible que Jacoubet a continué à

faire dessiner son plan puisqu'il était déjà en cours et réalisé *sous les bonnes auspices* de l'ancien préfet.

La soustraction de plans ou l'échec du projet de Chabrol.

Le projet de grand plan de Paris, initié par Chabrol en 1824, avait rapidement été divisé en 48 parties pour accélérer sa réalisation. Chacune des 48 parties avait été subdivisée en 24 plus petites, un nombre variable de ces parties, au 1/144^{ème}⁴³ étant alors assigné à chaque membre du bureau des plans. Sur le même modèle que les cadastres de Vasserot, Chabrol avait ajouté un niveau entre le grand plan et ceux des rues. Ainsi, un plan d'ensemble de chacune des 48 parties devait être réalisé en parallèle des plans des rues.

Les ordres de Chabrol ne sont suivis d'aucun effet puisque, le 12 juillet 1833, le nouveau préfet Rambuteau, constatant que le plan est à peine commencé, crée le poste de *chef de bureau des plans* qui se voit chargé de faire exécuter le plan et de répartir les crédits alloués au bureau dans ce but. Pourtant, cette mesure n'a aucun effet : si l'atlas de Jacobet progresse à grands pas, le plan officiel de l'Hôtel de Ville reste au point mort. Plusieurs témoins de la préfecture rapporteront que le bureau des plans se trouvait alors *dans un état de désordre tel* que tous les travaux menés étaient inutilisables ou jamais effectués (Gazette des Tribunaux, 1842c). On peut cerner quelques unes des causes de ce désordre.

Premièrement, le travail à la tâche encourage largement les géomètres et dessinateurs à privilégier la quantité sur la qualité. Les plans sont alors inexacts, incomplets ou faux, voire produits plusieurs fois. Le bureau met en place très tôt (peut-être dès 1823) un système de vérification censé empêcher ces malfaçons. Ainsi, lorsqu'un employé réalise un travail, il le présente à deux agents vérificateurs qui le contrôlent et, s'il est correct, le signent. Il suffit ensuite d'aller voir le responsable des paiements pour obtenir son dû en échange du plan. En 1827, Jacobet devient vérificateur en même temps que sous-chef, rejoint quelques années après par Hourdequin. Nul plan ne peut être payé sans leurs accords. Si ce système semble efficace, il échouera tout de même : en 1832, Morin imitera les signatures de Jacobet et Hourdequin obtiendra un paiement de 1512 francs. Il ne sera démasqué qu'en 1834 lorsqu'il tentera de nouveau d'obtenir paiement d'un travail déjà effectué pour la somme de 1883 francs.

Deuxièmement, plusieurs employés (Solet, Morin et Boutet) emportent chez eux plusieurs plans entre 1833 et 1838. Arguant qu'ils leurs servent à dessiner le plan d'alignement, certains les conserveront même après leur départ de la préfecture. Ainsi Solet, qui démissionne en 1836, conservera les plans du Champs de Mars, de la rue de la Verrerie et des Lombards que les enquêteurs retrouveront à son domicile en 1842. Les soustractions sont telles qu'entre 1823 et 1841, dates de deux inventaires des plans possédés par le bureau⁴⁴, plus de 500 plans - soit plus de 4% du total des

43. Il s'agit de la même échelle que les plans des rues de Verniquet, signifiant peut-être qu'il s'agissait d'une mise à jour de ces plans, dans le but de suivre le même procédé que celui qui avait été adopté en 1783

44. Un troisième est fait en 1834, mais les enquêteurs le jugent inutilisable car très incomplet

plans possédés- disparaissent.

Le travail est si inefficace qu'en 1841, un seul 48^{ème} du plan général d'alignements est terminé, montré lors du procès sous la forme d'un *grand plan de 6 pieds carrés*⁴⁵ sur lequel sont tracés en jaune tous les projets d'alignements et de percements.

On peut se demander alors quel statut occupe l'atlas de Jacoubet, qui est alors le seul plan détaillé et général de Paris, le plan en 48 parties étant un échec. La réponse est donnée par Jacoubet lui même, qui dans son interrogatoire précise que son plan était utilisé par les employés du bureau qui en faisaient des calques. Cela signifie que, pour réaliser les plans de chacune des 48 parties, les dessinateurs prenaient appui sur le tracé de l'atlas de Jacoubet. Celui-ci se justifie d'ailleurs d'avoir employé le personnel de la préfecture pour son propre atlas en invoquant l'utilité qu'il avait pour elle. Ainsi, bien que le l'atlas général n'ait jamais eu le titre de plan officiel, il sera de fait utilisé comme tel par le bureau des plans de Paris. Finalement, Jacoubet sera parvenu à la fois à faire de son travail une référence dans la sphère des géographes privés, mais également au sein même de l'Hôtel de Ville.

Le bureau des plans et la spéculation immobilière

Le dernier point du procès sur lequel nous allons nous attarder est celui des relations entretenues par certains employés de la préfecture et des entrepreneurs et agents d'affaires privés. L'essentiel des opérations spéculatives menées au sein du bureau des plans reposent sur une mécanique simple. En effet, lorsqu'un projet d'alignement ou de percement doit mener à la destruction de maisons et au retranchement de terrains, la Ville rachète ces terrains à l'issue de négociations avec chaque propriétaire. Le plus souvent, la transaction est très avantageuse pour le propriétaire. Bien conscient des enjeux, celui-ci revend le terrain à un prix supérieur à sa valeur réelle. Si le terrain est construit, il peut également obtenir le droit de revendre les matériaux issus de la démolition. Pour se prémunir d'une trop forte spéculation sur les terrains, la préfecture ne publie les plans d'expropriation que peu de temps avant le début des négociations et des travaux. De plus, ces plans ne sont pas diffusés publiquement mais consultables par les propriétaires au bureau des plans. Il s'agit alors pour les employés du bureau des plans, qui sont informés avant le grand public des différents projets, d'acheter à bas prix aux propriétaires ignorants les différents terrains expropriés ou reculés, puis de revendre ces mêmes terrains à la préfecture à l'aide de prête-noms conciliants ou intéressés.

Il s'agit cependant d'actes isolés, et non d'une corruption qui atteindrait l'ensemble de la préfecture. En particulier, nous nous attarderons sur les agissements de Hourdequin, alors chef de la Grande Voirie de Paris, à l'origine en 1843 d'une réorganisation du bureau des plans annonciatrice des changements adoptés par Haussmann 10 ans plus tard. Sur les 13 faits de corruption reprochés au chef de la Grande Voirie, 9 seront effectivement retenus (Gazette des Tribunaux, 1842b, pp. 3-4). Nous ne rapporterons ici que les 3 plus importants d'entre eux, qui suffisent à donner un aperçu révélateur du fonctionnement d'un des organes les plus importants de la préfecture de la Seine.

45. Soit environ 2 mètres carrés.

La première affaire renseigne sur la relation entretenue par Hourdequin avec l'entrepreneur en travaux publics Dubrugeaud dans un objectif d'enrichissement personnel. En juillet 1836, le conseil municipal adopte le projet de la rue d'Arcole qui devait être créée grâce à l'alignement des rues Chevet Saint-Landry et Saint-Pierre aux Bœufs. La concession est donnée à l'entrepreneur Dubrugeaud, qui se voit allouer la somme de 460.000 francs -payée après travaux- pour l'ensemble des opérations ainsi que le droit de revente des matériaux issus de la démolition, à la condition que les travaux soient réalisés en un an seulement. Terminé dans les temps, le percement de la rue d'Arcole permet à Dubrugeaud de dégager un bénéfice de 279 204 francs grâce à la vente des matériaux et à plusieurs opérations de spéculation immobilière dans la rue voisine de Constantine. L'extrême rentabilité de cette opération n'aura été pourtant possible que grâce à Hourdequin, alors associé de l'entrepreneur. Sur la rue d'Arcole elle-même, le tribunal ne relèvera que le zèle étonnant avec lequel il soutiendra la candidature de Dubrugeaud, puis un acompte de 10.000 francs qui permettront de terminer les travaux à temps. Surtout, les opérations menées sur l'île de la Cité seront l'occasion pour eux d'acheter plusieurs propriétés dans la future rue de Constantine avant même que le projet ne soit rendu public. Ignorant la démolition future de leurs maisons, plusieurs propriétaires les vendirent à bas prix à l'entrepreneur qui, informé des alignements et démolitions prévus par Hourdequin, pu les revendre à la ville avec une marge importante. En remerciement des services rendus par Hourdequin, son associé lui offrira finalement la somme de 17.000 francs.

La seconde affaire dévoile une corruption encore plus flagrante. En 1840, le baron Millin de Grandmaison, propriétaire de plusieurs immeubles sur la place de la Madeleine, avait prévu d'aligner ceux-ci, mais également de faire percer une nouvelle rue entre la place et la rue d'Anjou dans le but de revendre les terrains ainsi dégagés. Il soumet alors le projet au conseil municipal, ajoutant également l'abandon d'une maison d'une valeur de 40.000 francs se trouvant sur le tracé de la rue, en échange de l'acquisition d'une maison située au n° 16 rue de la Madeleine appartenant à la Ville et de la remise d'une soulte de 150.000 francs⁴⁶. Le conseil municipal puis le préfet refusant de céder plus de 100.000 francs⁴⁷, le dossier reste bloqué à la préfecture jusqu'à ce que M. de Grandmaison ne prenne contact avec Hourdequin. Grâce à un rapport élogieux remis au conseil municipal le 28 juillet 1841, ce dernier obtient l'accord de la Ville pour une somme réduite à 125.000 francs et réclame à Grandmaison la somme de 25.000 francs pour ses bons services. La lettre suivante, daté du 20 juin 1841, témoigne de l'accord entre eux :

Je vous envoie ma demande au préfet , mon cher Hourdequin, en vous priant de vous occuper de suite de cette affaire ; je prends l'engagement, si l'on m'accorde les 125 000 francs de soulte que je demande, de vous remettre sur cette somme 25 000 francs. Tout à vous d'amitié. Signé : baron de Grandmaison.

46. Dans un contrat d'échange, la soulte désigne une somme d'argent qu'une des parties remet à l'autre pour compenser un déséquilibre entre la valeur des objets échangés.

47. Une note du préfet Rambuteau le 7 juin 1841 indique ainsi : *Borner à 100 000 francs ; et s'il s'y refuse, ajourner l'affaire*

Le dernier exemple donne cette fois un aperçu d'actes de spéculation orchestrés directement par Hourdequin et rendus possibles par sa situation à la préfecture. Ainsi en 1833 il existait un projet d'allongement de la rue de l'Aiguillerie (aujourd'hui remplacée par la partie de la rue des Lombards entre la rue Sainte-Opportune et la rue Saint-Denis) jusqu'à la rue des Fourreurs (aujourd'hui rue des Halles). Connaissant le projet à l'avance, Hourdequin achète en août 1834 pour 36.100 francs la propriété située au n° 6 rue des Fourreurs⁴⁸ sous le nom de *Laurent Huet*. Le 12 février 1835, Hourdequin -sous le nom de Huet- et son voisin possédant les maisons 14 et 22 rue des Fourreurs⁴⁹, soumettent tous deux l'achat de leurs maisons à la Ville. Cette fois, le chef de la Grande Voirie use de sa position officielle et appuie fortement les deux demandes qui se trouvent acceptées par le conseil municipal dès le 6 mars 1835. La vente est consentie le 8 mai et finalement les maisons sont vendues le 4 septembre. Ainsi, les deux propriétaires auront atteint leurs objectifs en moins d'un an. La vente de la maison et des matériaux rapporteront 48.274 francs, formant donc un bénéfice de 12.174 francs. Son voisin, nommé Cady, obtiendra 80.000 francs pour la vente de sa maison, et 6.000 francs pour les matériaux. En remerciement, il versera 10.000 francs à Hourdequin.

Ainsi donc, Hourdequin profitera sa situation privilégiée à l'Hôtel de Ville pour s'associer avec divers acteurs privés, agents d'affaires mais aussi simples propriétaires profitant de l'occasion. Des trois affaires présentées, il retirera un bénéfice de 64.174 francs dont plus de 80% seront formés de "pots-de-vins". Pourtant, le jury ne le déclarera coupable que de l'opération menée avec Millin de Grandmaison, jugeant la plupart des autres affaires *peu morales* mais toutefois légales. Quant à Jacoubet, il apparaît totalement étranger aux diverses corruptions qui agitent le bureau des plans, bien qu'il connaisse plusieurs agents d'affaires mêlés au scandale⁵⁰. Pourtant, lorsque Eugène Deschamps, conservateur du plan et futur chef du service des plans sous Haussmann, est interrogé par le procureur sur l'éventuelle utilité de l'atlas de Jacoubet -qui indique tous les projets de voirie- pour les spéculateurs parisiens, celui-ci répondra : *Du jour où l'administration a autorisé un pareil plan, les spéculateurs ont pu avoir tous les renseignements possibles sans venir à nos bureaux. Cet Atlas leur offre même des renseignements plus complets, puisqu'il contient le plan de tous les îlots réunis.*

Si Jacoubet semble s'être tenu à l'écart de ces pratiques, il aura tout du moins créé un outil dont l'utilité pour les spéculateurs immobiliers ne semble pas avoir échappé au tribunal.

2.6.5 Notes et remarques

Sur l'atlas Général de Jacoubet et son devenir au bureau des plans

48. Il s'agit d'une propriété ayant issue *sur l'une et sur l'autre*. Ici le cadastre de Vasserot et l'atlas de Jacoubet se contredisent, la seule maison correspondant à la description, correctement numérotée sur Vasserot, est numérotée 8 sur l'atlas général.

49. Le n° 22, visible sur Vasserot, correspond à la maison dont l'issue est sur la rue de l'Aiguillerie, opposée à celle de Hourdequin. Le numéro 14 n'est pas indiqué, et se trouve à l'opposé de la rue sur Vasserot.

50. Notamment Denis-Laurent-Aubin Jaloureau, agent d'affaire de Hourdequin.

Si l'œuvre de Jacobet n'atteindra jamais la renommée de l'atlas de Verniquet, il n'en restera pas moins l'un des plans les plus importants de la première moitié du XIX^e siècle pour sa précision et sa richesse d'informations. Il permettra à son auteur de remporter la médaille d'argent à l'exposition des produits de l'industrie française en 1834, de se voir proposer une place à la chambre des Pairs et finalement d'obtenir une telle notoriété publique qu'il deviendra un expert reconnu de la topographie parisienne.

Il est difficile de savoir si l'atlas a été largement utilisé par la préfecture de la Seine. En effet, si un plan d'alignements du quartier des Halles daté de 1848 est basé sur l'atlas de Jacobet - il en est d'ailleurs un des auteurs -, rien ne permet d'affirmer qu'il soit directement lié à la Grande Voirie. En effet, le projet de réorganisation des Halles avait alors fait naître divers projets émanant d'architectes et géographes privés. En novembre 1842, le chef de la division de la Voirie Planson annonce que le plan de Chabrol, dont seul un quarante-huitième était fait en 1841, avait été *repris avec ardeur* en juillet 1842 et qu'il était désormais *fort avancé*. Cependant, lorsque Haussmann lancera la production du nouveau plan de Paris en 1856, la préfecture ne semble pas disposer d'un plan d'ensemble de Paris. Sans doute la réalisation du plan général annoncée par Planson n'était-elle pas aussi rapide qu'il le souhaitait.

De plus, le plan de Verniquet avait été utilisé par les services officiels au début du XIX^e siècle car aucun autre plan n'était aussi géométriquement exact. Au contraire, lorsque Jacobet publie son atlas le cadastre parisien est en grande partie terminé, le plan en 48 quartiers de Vasserot a été fini en 1821, celui par îlots est bien avancé et les plans des maisons seront terminés en 1836. Ainsi l'atlas n'est-il pas le seul plan général en lice. Les plans de percements des rues contemporains de l'atlas de Jacobet sont d'ailleurs tracés sur le plan parcellaire de Vasserot, bien plus adapté à des opérations impliquant expropriations et reculements. Jacobet n'atteindra donc qu'en partie son objectif : si son atlas est un vrai produit de l'administration de la Seine, sans laquelle il n'aurait jamais pu voir le jour, il ne sera jamais adopté officiellement par la préfecture, et s'il a aidé à la création du plan de Chabrol il ne semble ne l'avoir jamais remplacé.

Sur le bureau des plans

Au travers de l'affaire Hourdequin, nous avons pu avoir un aperçu des dysfonctionnement des services officiels destinés à gérer la voirie parisienne. Cependant la corruption à l'œuvre au bureau des plans n'y est pas circonscrite et d'autres affaires éclabousseront la préfecture. Ainsi le 7 décembre 1843, Trémery, ingénieur des Ponts-et-Chaussées et inspecteur général des carrières sous la Seine est contraint à la démission lorsque la justice découvre l'existence de 16 employés fictifs supposés travailler dans les carrières et dont les salaires sont détournés pour l'enrichissement personnel des dirigeants des carrières⁵¹. Pourtant la justice se montrera étonnamment clément : Trémery ne sera jamais poursuivi. Quant aux protagonistes de l'affaire Hourdequin seuls trois d'entre eux seront condamnés, Solet et Boutet étant acquittés. Si Hourdequin et Philidor se voient sanctionnés de quatre et trois ans de

51. En plus de Trémery, le chef Poirier-Saint-Brice sera placé en retraite.

prison pour corruption, la plupart des faits de corruption pourtant jugés vrais par le tribunal ne seront pas retenus. Ainsi, sur neuf faits qui lui sont imputés, Hourdequin ne se verra condamné que pour ses rapports avec Millin de Grandmaison. Quant à Morin, il sera finalement puni de trois ans de prison pour faux et usage de faux. Pour tous les accusés, le jury ne considérera pas la soustraction de plans. Pour conclure sur l'affaire Hourdequin, laissons le mot de la fin au journaliste de la *Gazette des Tribunaux* qui nous a permis de retracer l'histoire du plan de Jacoubet :

Nous n'avons pas à nous expliquer sur la décision qui vient d'être rendue (Le jugement final du 19 novembre 1842). Mais il y a dans ce procès une autre question que celle soumise au jury, et il ne faut pas que tout se termine avec son verdict. Nous ne voulons rien exagérer, rien amoindrir ; nous ne voulons pas rechercher si les condamnés sont les seuls coupables, ou s'il n'en manquait pas d'autres à côté d'eux sur les bancs de la Cour d'assises : ce sont là des questions qu'il ne nous appartient pas de résoudre, mais que les révélations du débat commandent à la justice d'étudier sérieusement.

...

Un tel état de chose appelle à un remède énergique et prompt : il faut que l'opinion publique soit rassurée ; il faut que l'administration elle-même soit mise à l'abri de cette espèce de suspicion légitime sous le coup de laquelle elle est placée. Car, c'est là une des tendances de l'opinion, d'exagérer le mal et de le généraliser. Sans

doute de telles questions sont graves et difficiles. Mais sans toucher imprudemment à l'ensemble du mécanisme administratif, on peut se demander si le contrôle et la surveillance sont ce qu'ils doivent être ; si les garanties de savoir et de moralité sont toujours sérieusement consultées à l'entrée de chacun dans la carrière administrative ; si enfin la question des salaires est tranchée de façon à placer le fonctionnaire au-dessus du besoin et de la corruption, ou si au contraire elle n'est pas dominée par ce besoin de créer incessamment des emplois inutiles et de pourvoir aux nécessités de la faveur.

Il semble que la Gazette ait été entendue puisque cette affaire causera la première grande réorganisation de la voirie parisienne, précédant celle opérée par Haussmann quelques années plus tard. En effet, le préfet Rambuteau, après avoir supprimé momentanément le bureau des plans, mettra en place une nouvelle hiérarchie avec l'objectif clair d'empêcher toute perte de plan et toute corruption. Ainsi le bureau sera désormais soumis à l'autorité directe de la Grande Voirie, divisé en quatre sections elles-mêmes placées sous l'autorité d'un géomètre en chef. Le poste de conservateur des plans, occupé depuis 1834 par Eugène Deschamps sera également profondément modifié. Indépendant des sections et placé directement sous l'autorité du géomètre en chef, le conservateur est désormais le gardien des plans d'alignements, dont l'accès n'est plus autorisé que sur dérogation du préfet. Cette organisation, visible dans la figure 2.5, restera la même jusqu'en 1857.

Sur Jacoubet

Jacoubet quant à lui continuera à éditer son atlas, d'abord en grand format puis sous sa forme réduite au 1/10.000^e et finalement l'étendra dans ses nouvelles limites en 1860 (Jacoubet, 1860). Nous ne savons pas s'il a conservé son poste au bureau des plans après la réorganisation de 1842 mais il est certain qu'il n'était plus à la préfecture après 1852. En effet, après avoir participé à la révolution de 1848 et s'être opposé au coup d'état de Napoléon III, il sera condamné à l'exil à Londres le 7 mars 1852⁵². Cette peine sera pourtant annulée et il se verra interné⁵³ dans la propriété de sa famille à Montesquieu-Volvestre près de Toulouse. Pourtant, c'est sans hésitation que le comte Siméon fera appel à lui en 1853 lorsqu'il formera la commission des embellissements de Paris voulue par Napoléon III. C'est en tant que cartographe et spécialiste de Paris que Siméon louera Jacoubet dans une lettre qu'il adresse à Persigny le 2 septembre 1853 (Casselle, 2000). Celle-ci renseigne d'ailleurs aussi bien que le procès sur un personnage dont le talent de commerçant n'avait rien à envier à celui de cartographe : non seulement Napoléon III *connaît Jacoubet*, pourtant républicain et condamné à l'internement, mais surtout celui-ci tente de vendre à l'empereur la gravure de son plan pour chaque rue à percer pour 1200 francs par rue, soit 4800 francs au total⁵⁴.

Hypothèse de travail

On remarquera que dans l'ensemble de nos développements autour de l'atlas général, nous n'avons jamais séparé clairement les phases de levé topographique des planches des phases de gravure. À l'inverse même nous avons considéré, tout comme Jacoubet l'annonce en 1842, que les minutes de son atlas étaient terminées en 1825. Notre choix s'explique pour deux raisons.

D'abord, nous avons montré que la version originale de l'atlas publiée en 1836 comportait de nombreuses corrections effectuées depuis la gravure des premières planches en 1827. Contrairement à l'atlas de Verniquet où le plan a d'abord été entièrement levé puis gravé une fois terminé, cette fois ce processus a été appliqué plusieurs fois sur des parties de l'atlas, à divers moments.

Ensuite, nous devons rappeler que la réalisation de l'atlas de Jacoubet suit l'édition de la version réduite et mise au net du plan de Verniquet en 1825, puisque le prospectus annonçant les premières livraisons date de 1826. Nous avons également fait remarquer précédemment que malgré les conclusions du procès intenté par M^{me} de Chavagnac à Jacoubet, il paraît difficile qu'un plan de l'ampleur de celui-ci ait été entièrement levé et dessiné par une seule personne (les quatre autres n'interviendront qu'après 1825) en une année. Notons également que le tribunal conclura non pas que l'atlas est entièrement nouveau, mais qu'il est suffisamment différent

52. On lit en effet dans La Presse du 7 mars 1852 qu'il *vient d'être délivré des passeports aux personnes [] qui se sont engagées à quitter Paris dans les 24 heures. Pour Londres : Jacoubet, architecte.*

53. Assigné à résidence

54. *M. Jacoubet, architecte, chargé autrefois du bureau des plans de la ville, possède tous les éléments de ce travail. Il s'engagerait, moyennant un prix à forfait de 1200 francs par rue à ouvrir, à fournir, en un mois, ces documents à la Commission. Pour quatre rues, ce serait une dépense de 4800 francs, qu'il vous serait facile d'imputer sur les fonds du ministère.* (Casselle, 2000)

du plan de 1825 par son contenu et sa triangulation pour être considéré comme une oeuvre originale, laissant ouverte l'hypothèse d'une filiation entre les deux atlas.

Nous posons donc l'hypothèse d'une filiation directe entre l'atlas général et le plan de Verniquet. En effet, la mise à jour du plan de Verniquet en 1825 a nécessité d'effectuer les relevés topographiques relatifs au rajout des nouvelles rues et places de la ville. Nous supposons donc que la triangulation effectuée par Jacoubet n'est pas tout à fait nouvelle, mais complète celle plus ancienne de Verniquet. Il utilisera en effet plusieurs fois d'autres plans pour compléter les informations de son atlas. Par exemples, les bâtiments en coupe qu'il représente sont issus de l'atlas des 48 quartiers de Vasserot. Si cette hypothèse se vérifie nous pourrions alors considérer l'atlas de Jacoubet comme l'héritier direct du plan de Verniquet pour le XIX^e siècle.

2.7 Conclusion

Au cours de ce chapitre, nous avons présenté les différents plans que nous utilisons dans leur contexte, à la fois pour montrer les spécificités de chacun et pour donner l'intuition des liens de filiation les unissent, ainsi que pour identifier les temps valides de chacun. En effet, il apparaît que ces différents plans et atlas, pourtant d'origines et de formes très variées, sont tous liés les uns aux autres. Ce lien peut être direct comme entre le plan de Maire, celui de Delagrive et l'atlas de Verniquet, ou encore entre celui des 48 quartiers de Vasserot et celui de Verniquet, où les plans les plus anciens sont utilisés comme fonds de cartes ou sont simplement mis à jour. Il peut être également indirect, parce qu'un même contexte les lie ou que la réalisation d'un plan découle de l'existence -ou plutôt de l'obsolescence- d'un autre, comme c'est le cas pour le plan de Haussmann avec celui de Jacoubet ou bien encore celui de Jacoubet avec Verniquet et Vasserot.

Lorsqu'elle est directe, la filiation entre deux plans est intéressante à considérer pour leur utilisation au sein d'un SIG. En particulier, les éléments suivants sont à considérer :

- Lorsque deux plans sont liés, les déformations géométriques du premier peuvent être corrélées à celles du second. Dès lors, il peut être préférable d'utiliser le premier comme référentiel pour référencer le second, plutôt qu'un référentiel absolu -souvent un plan récent-.
- Un lien direct peut permettre d'expliquer d'éventuelles incohérences dans le second plan, soit parce qu'il s'agit d'oublis, soit en raison du choix de son auteur (Maire par exemple figure des bâtiments pourtant détruits après 1803).
- Concernant le contenu des plans, un héritage direct entre deux plans indique une proximité sur les spécifications de saisie des objets géographiques et leur représentation cartographique lorsqu'ils sont communs (c'est par exemple le cas des bâtiments de l'atlas de Jacoubet, provenant du plan des 48 quartiers de Vasserot). Par conséquent, les spécifications des données extraites de ces plans par vectorisation doivent également être communes.

Lorsqu'elles sont indirectes, les filiations entre plans permettent de mieux comprendre les particularités d'un plan ou bien sa création. C'est par exemple le cas de l'atlas de Jacoubet qui n'aurait sans doute pas vu le jour sans l'absence de mise en place d'un remplaçant du plan de Verniquet par la préfecture. la puissance financière Au travers de ce chapitre, nous avons ainsi recomposé les filiations, résumées dans la figure 2.6. On voit donc que les plans sont largement liés entre eux, en particulier jusqu'au milieu du XIX^e siècle et qu'ils sont majoritairement affiliés à l'atlas de Verniquet. De plus, en cent ans, seules deux nouvelles triangulations véritables ont été effectuées, si l'on considère que celle de Jacoubet est au mieux partielle. Ceci renseigne d'autant plus sur le caractère exceptionnel du travail de Verniquet : la seule triangulation qui suivra sera l'œuvre d'un service entier de la préfecture de la Seine.

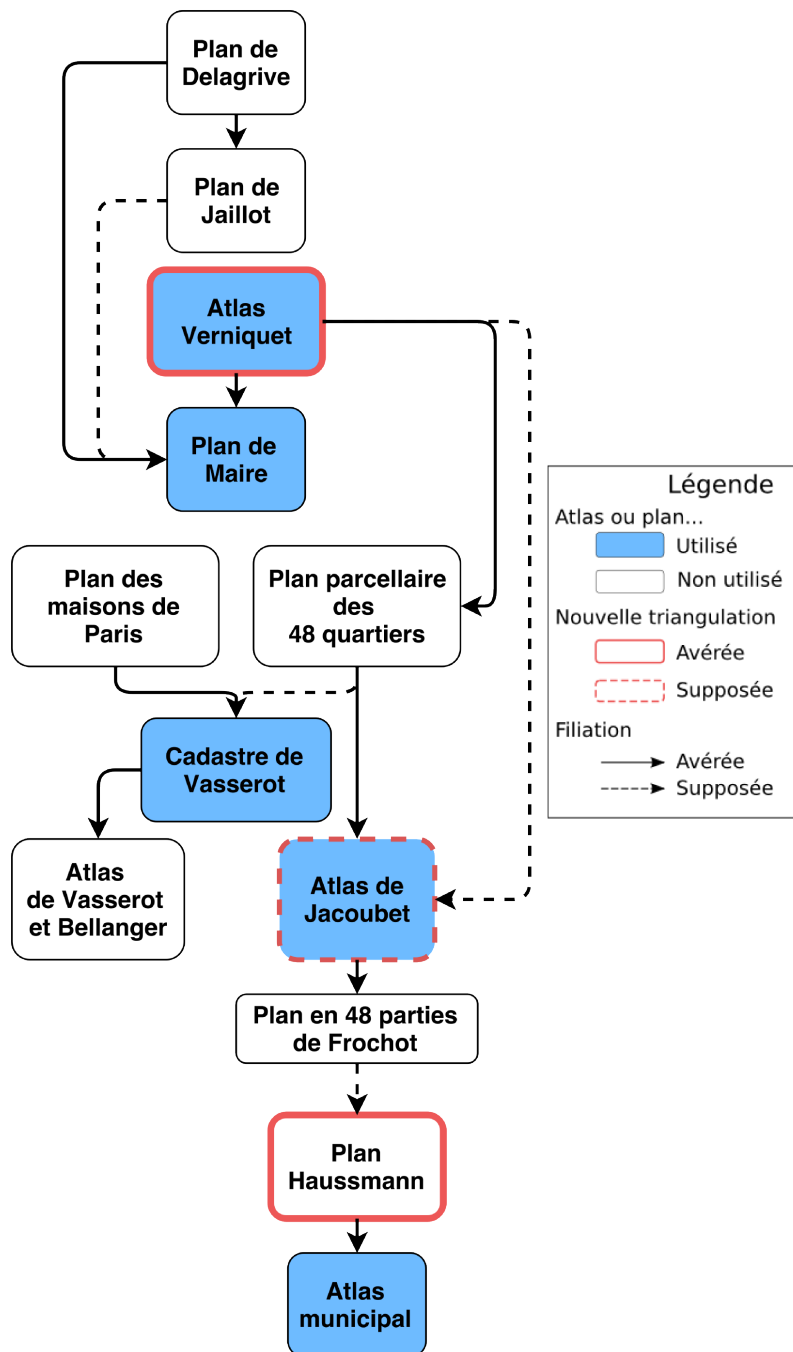


FIGURE 2.6 – Filiations identifiées entre les plans de Paris utilisés dans la thèse. Certains plans non utilisés mais cités dans le chapitre sont également inclus car ils font partie de cette généalogie

Deuxième partie

Construction de graphes spatio-temporels à partir de sources cartographiques anciennes

Chapitre 3

Géoréférencement des plans de Paris

Objectifs :

- Proposer une approche permettant de géoréférencer les atlas parisiens de façon précise.
- Proposer une approche visant à analyser la qualité géométrique des thèmes d'un plan topographique.
- Produire un corpus de données géographiques vectorielles à partir des atlas géoréférencés.

Sommaire

3.1	Construire une carte topographique au XIX^e siècle	114
3.1.1	Localiser un point sur la surface terrestre	115
3.1.2	Créer une carte topographique aux XVIII ^e et XIX ^e siècles	118
3.2	Géoréférencement : principe de la méthode usuelle	120
3.2.1	Types de géoréférencements	122
3.2.2	Principe détaillé du géoréférencement indirect	124
3.2.3	Types de fonctions de transformation dans le cas indirect	127
3.3	Évaluation de la qualité d'un géoréférencement indirect	129
3.3.1	Exactitude d'un géoréférencement	129
3.3.2	Qualité des points d'amer	130
3.4	Géoréférencement des plans de Paris	133
3.4.1	Choix du système de projection des plans topographiques anciens de Paris	134
3.4.2	Choix du système de projection du référentiel géographique cible	137
3.4.3	Référentiels géographiques actuels utilisables pour le placement des points d'amer	137
3.4.4	Choix des points d'amer et des modèles de transformation spatiale	142
3.4.5	Conclusion sur l'approche de géoréférencement proposée	145
3.5	Cas applicatif : l'atlas de Verniquet	147
3.5.1	Construction du carroyage	147
3.5.2	Placement des points d'amer sur les différents thèmes cartographiques et géodésiques	149
3.5.3	Choix du type de points d'amer et d'une transformation spatiale	150
3.5.4	Vérification du placement de la perpendiculaire et de l'approximation du système de projection de l'atlas par le Lambert I NTF	155
3.5.5	Analyse des distorsions des thèmes "bâti" et "rues" de l'atlas . . .	166
3.5.6	Conclusion et géoréférencement des autres plans de Paris	171
3.6	Vectorisation des plans	171
3.6.1	Classe d'objets vectorisés et spécifications de saisie	172
3.6.2	Résultats de la vectorisation	173
3.7	Conclusion	174

Au travers du premier chapitre, nous avons vu que la reconstitution des états d'un objet géohistorique pouvait être effectuée en deux temps : d'abord en extrayant de sources anciennes des représentations de cet objet (le tracé d'une rue, le plan d'un bâtiment, etc.), puis en les liant pour reconstituer l'objet géohistorique sous-jacent. Ces liaisons reposent sur l'identification de persistances et de changements entre représentations consécutives. Les sources étant cartographiques, la composante spatiale est un élément fondamental de cette identification. Pour avoir une connaissance des propriétés spatiales des observations, deux étapes sont nécessaires : une première pour déterminer leurs positions les unes par rapport aux autres ou par rapport à un référentiel spatial et une seconde pour extraire leurs formes des sources cartographiques.

Le référencement des cartes pourrait être fait de façon relative en les mettant en correspondance avec une carte de référence à l'aide de méthodes d'appariement d'images. L'application de telles méthodes a par ailleurs été proposée dans le contexte de cartes historiques par Balletti (2000). Cependant, l'extraction d'observations géohistoriques permet d'envisager l'intégration de bases de données géographiques extérieures, anciennes ou récentes pour enrichir la base de données spatio-temporelle. Une seconde possibilité consiste à situer les différentes sources dans un référentiel commun, indépendant d'une source particulière. Celles-ci décrivant une portion de l'espace géographique, il est assez naturel de considérer ce dernier comme référentiel : il s'agit donc de géoréférencement. Les observations extraites de ces cartes géoréférencées sont localisées dans un référentiel géographique commun dès leur création, facilitant ainsi l'intégration de données géographiques externes.

Dans ce chapitre, nous nous intéressons dans un premier temps au géoréférencement de nos sources cartographiques anciennes, puis à la vectorisation de leur contenu. Pour ce faire, nous proposons une approche de géoréférencement adaptée pour le cas spécifique des plans topographiques anciens de Paris. Afin d'introduire les particularités de cette approche, nous présentons tout d'abord en section 3.1 les problématiques posées par le géoréférencement de plans dressés avec des techniques et des connaissances géodésiques particulières sur des référentiels récents. En section 3.2, nous présentons en détail le processus de géoréférencement ainsi que les notions liées à la qualité d'un géoréférencement 3.3, puis, en section 3.4 nous introduisons les adaptations que nous proposons pour le cas des plans de Paris. Cette proposition est par la suite testée et validée en l'appliquant à l'atlas de Verniquet (3.5). Par la suite, nous montrons comment cette approche nous permet d'effectuer une analyse quantitative de la qualité géométrique d'un plan. Ceci nous permet notamment de poser l'atlas de Verniquet comme un référentiel géographique géométriquement précis sur l'espace ancien de Paris. Après avoir appliqué cette approche aux autres plans de Paris, nous présentons en section (3.6) le processus de vectorisation nous ayant permis d'extraire un ensemble de données géographiques vectorielles pour chaque plan de Paris qui nous permettent de peupler notre modèle de suivi d'objets géohistoriques.

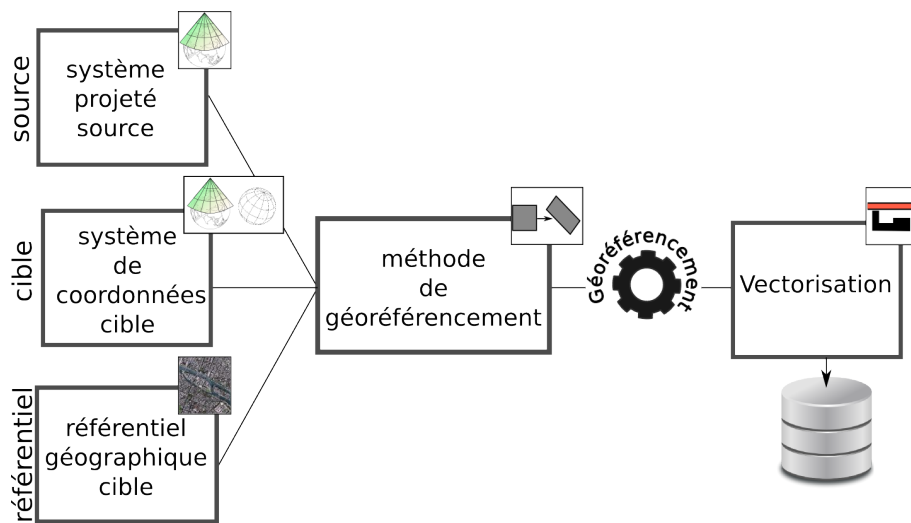


FIGURE 3.1 – Schéma du chapitre 3.

3.1 De la Terre à la carte : systèmes de coordonnées actuels et contemporains des atlas parisiens

Géoréférencer un plan, c'est attribuer à une représentation plate de l'espace des coordonnées dans un modèle de la surface terrestre. Différents modèles se sont succédés au cours du temps dans le but d'approximer le mieux possible la forme complexe de la planète. L'évolution rapide de l'astronomie à partir du XVII^e siècle a donné lieu aux premiers modèles mathématiques de la Terre sous la forme d'une sphère. Les travaux des Cassini et d'autres astronomes de l'Académie Royale des Sciences pour déterminer le méridien de Paris et surtout sa position sur le territoire français, nommé Méridienne de France, ont fixé -au moins pour la cartographie française- un ensemble de repères géodésiques pour les topographes et cartographes.

Cependant, les modèles de la Terre tout comme le tracé de la méridienne ont continué à évoluer. Aujourd'hui, ni le modèle sphérique du XVIII^e ni le méridien de Paris ne sont plus utilisés et les données géographiques sont localisées suivant des modèles plus complexes. La carte ancienne n'est donc pas une image simple, mais elle est elle-même associée à un modèle de la Terre qu'il faut prendre en compte. Celui-ci est toutefois inconnu ou difficile à déterminer avec précision, faute de connaissances sur la phase de levé topographique du plan. Généralement, le modèle de la carte ancienne est simplement ignoré et la carte considérée comme une simple image, ce qui peut entraîner des distorsions et des erreurs lors du géoréférencement. À l'inverse, connaître le support géodésique sur lequel s'appuie un plan permet d'éliminer les erreurs liées aux écarts entre modèles pour ne conserver que celles des seules opérations de levé.

Dans cette section, nous présentons la stratégie adoptée en ce sens pour les atlas de Paris. En particulier, nous montrons que certaines spécificités de la cartographie parisienne liées à sa proximité avec le tracé physique de la méridienne de France¹ nous permettent d'utiliser des systèmes de coordonnées actuels comme équivalents

1. Voir le paragraphe 3.1.2

de ceux utilisés au moment du levé des plans. De cette façon, nous minimisons les distorsions des plans au moment du géoréférencement dues aux écarts entre systèmes actuels et systèmes anciens.

Dans ce but, nous revenons d'abord sur la façon de modéliser la Terre et de localiser des objets géographiques à sa surface, puis nous aborderons rapidement les pratiques de levé topographiques des cartographes parisiens aux XVIII^e et XIX^e siècles. Enfin, nous présenterons le choix d'un système de coordonnées projetées adapté à nos atlas de Paris.

3.1.1 Localiser un point sur la surface terrestre

Figures de la Terre

Pour situer un point à la surface de la Terre de façon univoque, il faut d'abord disposer d'un modèle décrivant sa forme, généralement appelé *figure de la Terre*. Celle-ci est complexe et irrégulière, soumise à des dynamiques d'origine géologiques (séismes, tectonique des plaques, érosion, etc.) qui s'accumulent pour la distordre encore plus. Au cours de l'Histoire plusieurs modèles de complexité croissante ont été proposés. Depuis l'hypothèse de rotondité de la Terre vérifiée par Ératosthène et jusqu'au XVII^e siècle la planète a été approximée par une sphère (voir la figure 3.2a). La sphère de Picard, créée en 1670, est sans aucun doute l'une des plus connue. Les travaux de Huygens en 1673, puis ceux de Newton et Cassini montrent les limites du modèle sphérique, trop approximatif pour permettre des calculs de distances et d'angles précis. Les calculs des deux derniers scientifiques, pourtant contradictoires, prouvent que la forme géométrique qui approxime le mieux la Terre n'est pas une sphère mais un ellipsoïde de révolution (3.2b), c'est à dire une sphère aplatie aux pôles. À partir de ce moment, de nombreux modèles ellipsoïdaux apparaissent, certains d'entre eux restant encore aujourd'hui utilisés, en particulier pour la cartographie. De tels modèles simplifiés ne conviennent cependant pas pour des mesures de haute précision nécessaires dans des domaines tels que l'aéronautique. Ces considérations ont mené au dernier modèle existant : le géoïde (3.2c). Celui-ci représente la surface du potentiel de pesanteur terrestre coïncidant le mieux possible avec le niveau moyen des océans. Autrement dit, il s'agit d'une approximation de la Terre affectée par la répartition des masses de matière dans la croûte terrestre. Comme nous l'avons dit, l'ellipsoïde est préféré en cartographie en raison des difficultés à projeter sur la surface plane de la carte les irrégularités dues au géoïde. Afin de minimiser les erreurs liées à l'approximation ellipsoïdale, ces modèles sont souvent locaux et adaptés à des zones du géoïde. Ces ellipsoïdes locaux sont donc choisis de façon à minimiser l'écart entre leur surface et celle du géoïde sur une étendue donnée. Ainsi, pour la France, le modèle ellipsoïdal IAG-GRS80 est utilisé depuis 2000 en remplacement de l'ellipsoïde Clarke 1880 sur lequel étaient construits les systèmes de projection cartographiques français Lambert I à IV (Girres, 2012).

Systèmes géodésiques

L'intérêt d'une figure de la Terre est de pouvoir positionner des objets à sa surface. Un *système géodésique*, ou *datum* est un ensemble de constantes et de méthodes permettant de définir un repère affine qui assigne à chaque point de la Terre une

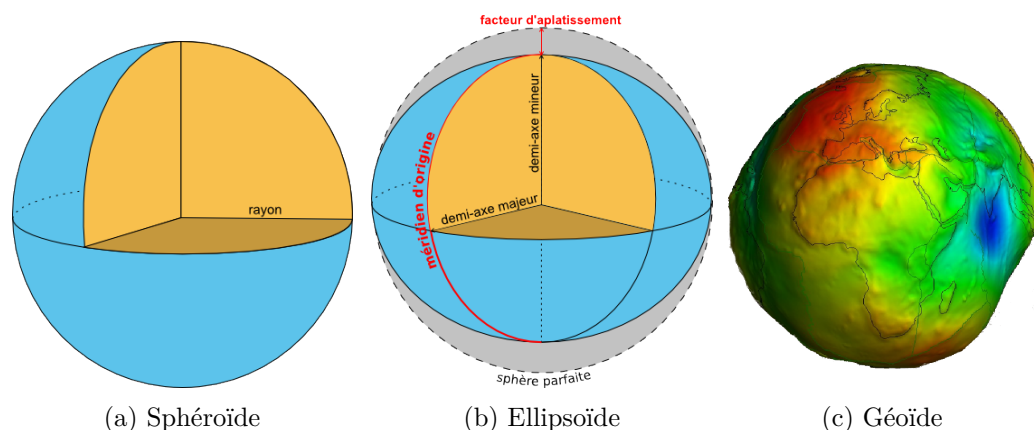


FIGURE 3.2 – Les trois figures de la Terre, ordonnées par qualité d’approximation croissante. L’aplatissement de l’ellipsoïde est largement exagéré dans cet exemple.

coordonnée unique². Un système géodésique se compose des éléments suivants :

- un ellipsoïde de référence défini par ses paramètres de demi-axes majeur et mineur, d’aplatissement et son écart au centre des masses de la Terre,
- un modèle de géoïde de référence permettant de mesurer des altitudes,
- un système de coordonnées géographiques en trois dimensions associé à un méridien d’origine sur l’ellipsoïde de référence,
- un réseau de points géodésiques qui sont des points de repère calculés avec une très grande précision dans le système de référence.

Un datum géodésique est associé à un *réseau géodésique*, un ensemble de points physiquement ancrés dans le sol permettant de mesurer les variations des positions des points de la surface (en raison de la tectonique des plaques, des changements d’altitude, etc.) et de corriger les mesures. Certains datum sont *locaux* car ils définissent en plus un ellipsoïde de révolution local destiné à amoindrir les approximations dues à l’ellipsoïde de référence. Dans ce cas, on associe toujours à un ellipsoïde local son positionnement par rapport au centre d’un système standard³. Les systèmes comportent généralement au moins un système de coordonnées planes dédié à la représentation cartographique.

Il existe un certain nombre de systèmes géodésiques standards, globaux ou locaux. En France, le système global officiel depuis le 1^{er} janvier 2001 est le *réseau géodésique français 1993* (RGF93), qui s’appuie lui-même sur le réseau géodésique européen ETRS89⁴. Il remplace en particulier le système local NTF⁵, construit de 1884 à 1991 par raffinements successifs d’une triangulation sur le territoire de la France métropolitaine.

Comme nous traitons de plans topographiques, nous ne nous attardons pas sur tous les constituants d’un datum, mais nous détaillons seulement les systèmes de coordonnées planes.

2. http://education.ign.fr/sites/all/files/geodesie_systemes.pdf

3. Actuellement, il s’agit du système WGS84, utilisé notamment par le GPS.

4. *European Terrestrial Reference System 1989*

5. *Nouvelle Triangulation de France*

Projections cartographiques

Nous l'avons dit, la cartographie impose de projeter dans un plan un espace pourtant irrégulier et approximativement ellipsoïdal. Or, il est impossible de développer un ellipsoïde en surface plane (Pumain et Béguin, 2010). La solution est alors de projeter des parties de cet ellipsoïde sur une surface qui est, elle, développable. La figure 3.3 illustre comment un cylindre peut servir de surface de projection. Ce procédé a cependant des inconvénients car il ne permet pas de conserver à la fois les angles et les surfaces, déformant de fait les objets géographiques. De plus, ces altérations augmentent à mesure que la zone à projeter s'agrandit et que l'on s'éloigne du parallèle tangent à la surface de projection (correspondant aux lignes rouges sur la figure 3.3). Pour cette raison, il est nécessaire de connaître parfaitement leurs effets, en particulier en ce qui concerne l'altération linéaire qui fausse le calcul des distances euclidiennes dans le système projeté⁶.



FIGURE 3.3 – Développement de la surface terrestre sur un cylindre. Le parallèle tangent est situé sur l'équateur (projection Mercator).

Il existe trois types de projections cartographiques, dont deux sont définies selon qu'elles conservent :

- les angles seulement ; on parle alors de projections conformes.
- les surfaces seulement ; il s'agit de projections équivalentes.

Une projection ne peut donc pas être conforme et équivalente, mais elle peut ne conserver ni angles ni surfaces : il s'agit alors de projection aphylactique.

Au delà de ces trois types de projections, la surface développable choisie influe également sur le résultat. Trois d'entre elles sont particulièrement usitées, le cylindre, le cône et le plan (on parle alors respectivement de projections cylindriques, coniques et azimutales). Il existe donc autant de catégories de projections qu'il est possible de combiner les surfaces développables et les types de projections. Par exemple, la projection Mercator, proposée en 1569, conserve les angles et utilise un cylindre pour passer du globe à la carte, il s'agit donc d'une projection cylindrique conforme.

De la Terre à la carte et son système de coordonnées

Un système géodésique muni d'une projection permet donc de passer de l'espace à la carte, c'est à dire attribuer à un objet situé à la surface de la Terre des coordonnées dans le plan. Ceci nécessite finalement de connaître les paramètres suivants :

6. L'altération linéaire se définit comme la variation des longueurs induite par le passage d'une figure de la Terre à une projection plane.

- un datum géodésique ;
- un méridien d'origine qui permet de construire un repère de référence dans lequel exprimer les coordonnées ;
- un type de projection qui permet de conserver certaines caractéristiques géométriques des objets.

Lorsque l'on souhaite replacer une carte ancienne dans un système de coordonnées géographiques, il faut effectuer une transformation de son système de coordonnées projetées propre vers un système de coordonnées géographiques connu. Il faut donc connaître les paramètres du système de projection de la carte ancienne, ou tout du moins les approximer de façon à minimiser les erreurs lors de la transformation. Lorsque l'échelle des cartes est grande, comme à Paris, on peut supposer que ces altérations sont négligeables par rapport aux distorsions de la carte résultant d'erreurs de levé et de l'altération de la carte elle-même.

Nous allons maintenant explorer la façon de construire une carte aux époques qui concernent nos plans de façon à estimer leur système de projection.

3.1.2 Créer une carte topographique aux XVIII^e et XIX^e siècles

En général, le système de projection d'une carte ancienne est peu connu parce que la carte finale ne comporte pas les indications géodésiques de son relevé et que les minutes des relevés topographiques se sont perdues. Ces informations sont d'autant moins connues que l'on recule dans le temps : avant le XVII^e en France, aucun système géodésique standard n'existe. Nos plans topographiques parisiens font cependant exception. En effet, ils apparaissent à une époque de révolution cartographique qui voit s'établir le premier système géodésique précis s'appuyant sur le méridien de Paris et la figure de la Terre sphérique dite *sphère de Picard*. Afin de bien comprendre en quoi consiste la création d'une carte topographique à l'échelle d'une ville aux époques qui nous concernent, nous proposons de faire ici un point sur la méthode géométrique adoptée par les cartographes et géomètres pour transcrire le territoire sur une carte : la triangulation. Nous verrons dans la section suivante que les particularités des plans topographiques de Paris permettent d'approximer leur système de projection.

La méridienne de France, socle des opérations de triangulation à Paris

Pour réaliser un levé topographique, il faut disposer d'un repère à partir duquel les coordonnées des points relevés seront exprimées. Le calcul de la méridienne de France (aujourd'hui appelé "méridien de Paris") permet de fournir un repère absolu sur lequel tous les cartographes et ingénieurs topographes peuvent s'appuyer, les coordonnées étant exprimées par rapport à cette méridienne⁷. Pour obtenir *des cartes géographiques de la France plus exactes que celles qui ont été faites jusqu'ici* (Martin et Virol, 2008, p.235), l'astronome Jean Picard est nommé par Colbert responsable de la réalisation d'un "Châssis général" de la France (Gallois, 1909, 291–295). Ce canevas géodésique repose sur deux éléments :

- le relevé de la méridienne de Paris,

7. Pour les latitudes, le repère est plus simple puisqu'il est fondé sur l'équateur.

- de la création d'un canevas de triangles couvrant toute la France s'appuyant sur cette méridienne ;

La méridienne s'ancre alors physiquement dans Paris avec la construction de l'Observatoire Royal, terminée en 1671, aligné sur le tracé de la méridienne et traversée par elle en son milieu. Le calcul de la méridienne de Paris allant de Dunkerque à Perpignan, issu du travail de Picard, puis de Jean-Dominique Cassini et Jacques Cassini, est terminée en 1718 (voir la figure 3.4a)⁸. Afin de définir un plan dans lequel le canevas triangulaire de la France peut être réalisé, cinq perpendiculaires à la méridienne sont ajoutées⁹, dont une passant par l'Observatoire Royal, visible dans la figure 3.4b

Le canevas de triangles couvrant la France est quant à lui réalisé par César-François Cassini de Thury à partir de 1756 pour la carte de même nom.



(a) Premier canevas et méridienne de France (Cassini de Thury, 1744b)

(b) Détail à Paris. Extrait de (Cassini de Thury, 1744a)

FIGURE 3.4 – Le premier canevas triangulé de la France, avec la méridienne et les cinq perpendiculaires (à gauche). La position de la méridienne et de sa perpendiculaire à Paris sont visible à droite : on les voit se croiser à l'Observatoire.

Triangulations géodésiques parisiennes

L'idée de la triangulation géodésique est de construire sur l'espace géographique un maillage triangulaire dont les longueurs et l'orientation sont connues par rapport à un repère (ici la méridienne) (Boutier, 2002). Le maillage est réalisé dans un espace plan de façon à n'utiliser qu'une trigonométrie euclidienne. La méridienne est alors approximée localement par un segment de droite nommé *base*, ce qui oblige à créer plusieurs triangulations locales s'il faut couvrir une grande portion du territoire.

La construction du maillage est un processus itératif, schématisé en figure 3.5. D'abord (point 1), une base AB de longueur et d'azimut connus est déterminée.

8. Cela signifie qu'un réseau de points matérialisés est géoréférencé.

9. Il ne s'agit pas de parallèles géodésiques.

Il s'agit généralement d'une portion de la méridienne. Un troisième point C correspondant à un repère visuel fixe (clocher, tour, colline, etc) est choisi et les angles α_1 et α_2 sont mesurés (point 2). On peut alors construire le triangle ABC, puis recommencer avec un point D en utilisant un coté du triangle ABC comme nouvelle base (point 3). À Paris, la proximité de repères physiques de la méridienne (l'Obser-

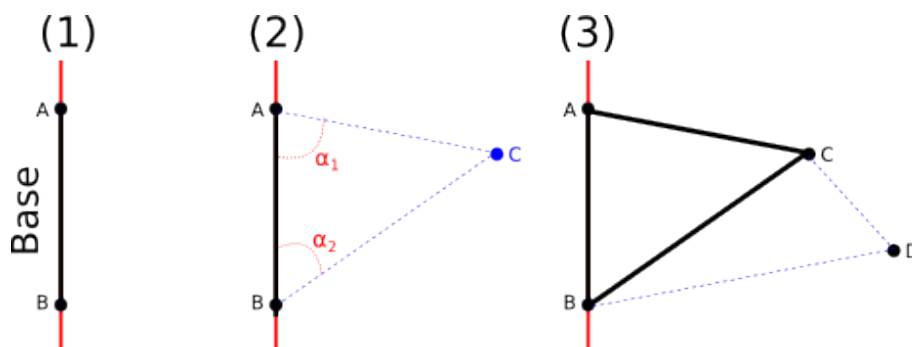


FIGURE 3.5 – Schéma simplifié des étapes d'une opération de triangulation géodésique.

vatoire Royal et deux mires placées aux extrémités nord et sud de Paris) offre aux cartographes la possibilité de créer des canevas de triangulation extrêmement complets. Souvent, ils s'appuient non seulement sur la méridienne mais également sur sa perpendiculaire passant à l'Observatoire Royal. La précision géométrique devenant une préoccupation majeure des cartographes de la fin du XVIII^e siècle, le canevas de triangulation est mis en avant et accompagne parfois même les atlas (c'est le cas pour les atlas de Verniquet et Jacoubet par exemple). Dans le cas de la cartographie parisienne, qui impose un niveau de granularité fin en descendant au niveau de la rue ou du bâtiment, d'autres types de relevés topographiques complètent la triangulation. En particulier, le levé des rues est effectué en traçant non plus des triangles mais des séries de segments formant des lignes brisées dont l'orientation de proche en proche est connue et rapportée à un triangle du canevas¹⁰. La figure 3.6 en présente un exemple tiré des minutes de l'atlas de Verniquet.

Tous nos atlas sont construits de cette façon, soit à partir d'une nouvelle triangulation, soit en réutilisant des canevas existants¹¹. Tous les relevés topographiques de nos atlas sont donc effectués dans un système de coordonnées projeté relatif à la méridienne de France.

3.2 Géoréférencement : principe de la méthode usuelle

Géoréférencer un plan topographique est une opération qui consiste à situer ce plan dans un repère géographique muni d'un système de coordonnées géographiques dans le but de croiser les informations qu'il contient avec celles d'autres représentations de l'espace (Balletti, 2006). En général, le plan lui-même possède un système

10. En topographie, il s'agit d'un *cheminement polygonal*. Les coordonnées de chaque point du chemin peuvent être calculés si on connaît les longueurs des segments et les angles de chaque paire de segments consécutifs.

11. Tout comme les cartes France entière de Cassini ou d'État Major

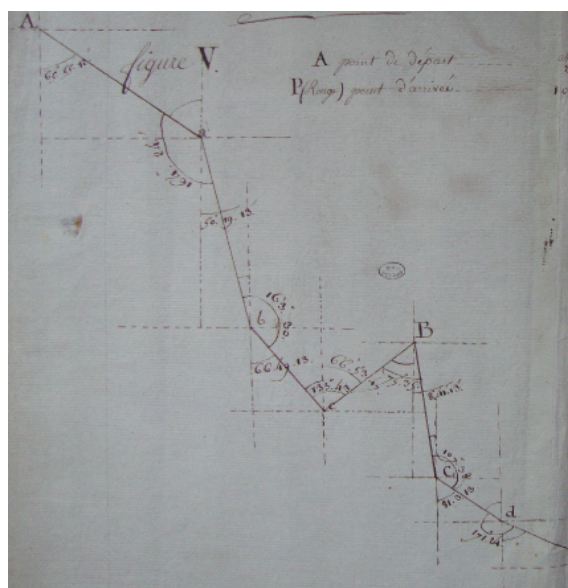


FIGURE 3.6 – Cheminement polygonal pour le relevé d’une rue. Extrait des minutes de l’atlas de Verniquet.

de coordonnées projetées. La figure 3.7 illustre le processus de géoréférencement de façon très générale. Un géoréférencement repose sur trois paramètres :

1. le système de coordonnées projetées de la carte source à géoréférencer,
2. une méthode de transformation,
3. un système de coordonnées cible (géographique ou projeté) dans lequel transformer la source.

Connaître le système de projection de la source est cependant optionnel. En effet, si on ne connaît pas le système source on utilise des points de repère pour projeter directement l’image dans le repère cible.

Dans la section précédente, nous avons déterminé le premier de ces trois points. Dans cette section, nous allons détailler le processus de transformation lui-même.

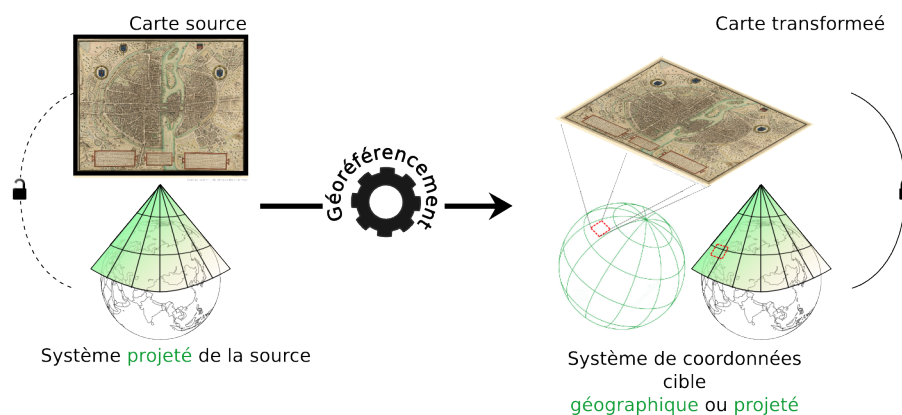


FIGURE 3.7 – Schéma global du géoréférencement

3.2.1 Types de géoréférencements

Une méthode particulière de recalage d'images

Le géoréférencement de cartes appartient au domaine de la géodésie et est lié aux méthodes de recalage d'images (ou mise en correspondance d'images), qui consistent à optimiser un critère de similarité entre deux images (Zhang, 1993, pp.11–12). Recaler des images entre elles est un problème général qui sert, selon les disciplines, des desseins variés : fusion d'informations, détection de changements, reconstruction 3D à partir d'images stéréoscopiques, reconnaissance de formes. Les méthodes de recalage sont classées en trois catégories, selon le critère utilisé (Brown, 1992). Les méthodes *géométriques* utilisent des critères mesurant la correspondance entre des primitives géométriques communes aux deux images. Elles peuvent être intrinsèques à l'image (points, contours, etc.) ou dépendantes des appareils d'acquisition (capteur, orientation de l'appareil, etc.). Une autre catégorie, fondée sur l'intensité, traite des images dans leur totalité en supposant qu'elles sont liées par une relation affine ou d'égalité : il s'agit des méthodes *iconiques*. Il existe enfin des méthodes hybrides associant des processus issus à la fois des méthodes géométriques et iconiques.

Quelque soit la méthode choisie, il est nécessaire de définir une **fonction de transformation** que l'on note $F : I_s \rightarrow I_c$ dont l'objectif est de maximiser la valeur d'un critère de similarité entre une image source I_s et une image cible I_c . Si cette fonction dépend du type de données à projeter, son objectif est constant : appliquée à l'image source, elle la projette dans le repère de l'image cible. Dans le cas du géoréférencement, les images sources et cibles ont la particularité de posséder des systèmes de coordonnées géographiques ou projetés.

Il existe deux approches du géoréférencement, complémentaires, nommées respectivement *directe* et *indirecte* s'appuyant sur les propriétés de l'image ou sur des primitives géométriques contenues dans celle-ci.

Le géoréférencement direct

Le géoréférencement direct est essentiellement utilisé en photogrammétrie pour situer des données de télédétection¹² à la surface de la Terre. Il s'agit d'un recalage dans lequel la transformation spatiale est directement calculée à partir de l'orientation de l'image par rapport au référentiel géographique, déterminée grâce aux paramètres internes et externes du capteur. Les premiers dépendent de la nature du capteur (position du point principal, focale, distorsions de la lentille), tandis que les seconds sont généralement calculés à l'aide d'une centrale inertielle et d'un GPS (Schaer *et al.*, 2003). Nous n'entrerons pas plus avant dans les détails, le géoréférencement direct n'étant pas le propos de notre travail. Il est cependant important de comprendre qu'une telle méthode implique deux choses. D'une part, elle n'est utilisable qu'à condition de connaître parfaitement les propriétés des outils ayant permis l'acquisition de l'image. D'autre part, elle suppose une maîtrise maximale des distorsions de l'image pour qu'elles soient rectifiées par la transformation.

Géoréférencement indirect

12. Images aériennes ou satellitaires, relevés RADAR, LiDAR ou laser, etc.

Il s'agit cette fois du procédé utilisé dans le cas où l'on ne connaît pas *a priori* les paramètres de la fonction de transformation F . Ceux-ci sont alors estimés grâce à l'identification d'éléments communs à l'image et à l'espace géographique. Le plus souvent il s'agit de points, désignés dans la littérature de différentes façons : *points de contrôle*, *de calage*, *d'amer* ou encore *d'appui* selon les domaines. On préférera ici le terme de **point d'amer**, qui désigne un objet fixe et remarquable. Lorsque l'on dispose de deux images I_1 et I_2 , on notera deux points d'amer correspondant à des éléments communs dans I_1 et I_2 par le couple (p, p') où p est un point dans I_1 et p' un point dans I_2 . La figure 3.8 présente un exemple de points d'amer placés sur l'Hôtel des Invalides à Paris. Le critère de similarité utilisé dépend d'une représentation de la Terre. La distance euclidienne entre les points d'amer est la plus souvent utilisée quand cette représentation est située dans le plan. Dans les autres cas, des distances géodésiques sont plus adaptées¹³. Notons également que des primitives géométriques plus complexes que des points peuvent être utilisées comme éléments communs. Par exemple, Baiocchi et Lelo (2005) s'appuient sur des structures linéaires pour géoréférencer une carte ancienne de Rome. De la même façon, Chryssa Papakosta et Vassilaki (2009) utilisent le tracé de routes et de cours d'eau pour géoréférencer d'anciennes vues aériennes de Grèce.

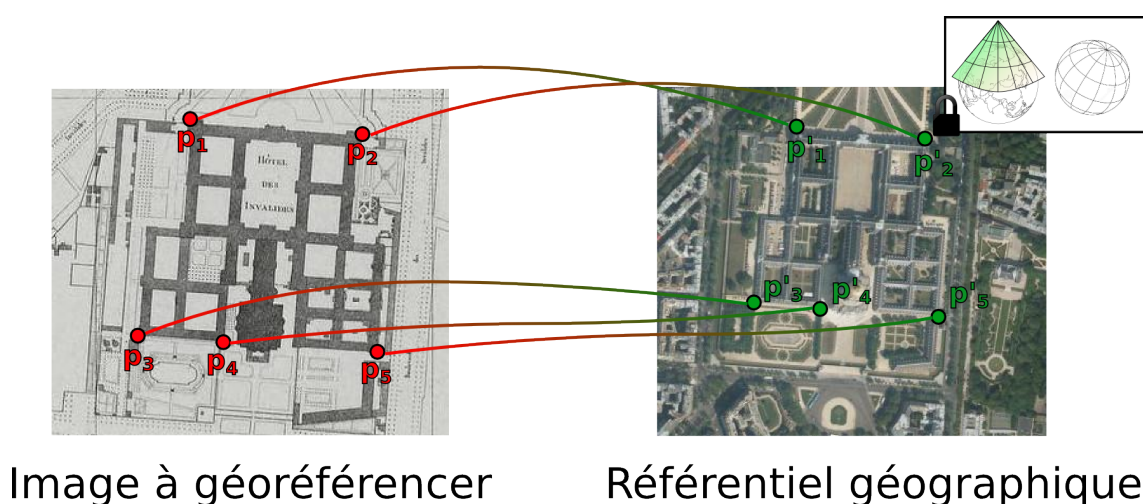


FIGURE 3.8 – Un exemple de placement de points d'amer

Le cas des cartes anciennes

Dans le cas où les images recalées sont des cartes anciennes, la méthode directe apparaît inapplicable. D'une part, le système de coordonnées projetées de la source -la carte ancienne- est généralement inconnu (voir la section 3.1.2). D'autre part, les cartes, proposant des représentations visuelles variées, sont peu adaptées à un géoréférencement direct s'appuyant sur des similarités entre propriétés des images. C'est pourquoi dans la suite de cette thèse, nous avons recours à la méthode indirecte fondée sur l'identification de points d'amer. Cependant, nous devons être conscient d'une limite importante de la méthode indirecte, en particulier pour des

13. La distance géodésique est une distance à la surface d'un ellipsoïde. Par commodité, elle est souvent approximée par la distance euclidienne.

cartes anciennes. Le placement de points d'amer sur un référentiel récent est parfois difficile en raison de la trop forte évolution du territoire qui empêche l'identification de points homologues. Or dans l'espace urbain en forte mutation qu'est Paris au tournant du XIX^e siècle, rien n'est plus incertain que la découverte de tels points, en particulier dans les zones périphériques. En réalité, la plupart des faubourgs de l'Ancien Régime disparaissent entièrement sous les nouveaux lotissements des années 1820-1830 ; la figure 3.9 le montre sur l'ancien espace du faubourg du Roule. On y voit très clairement la rupture provoquée par l'opération immobilière entre les plans de 1808 et 1836 qui épargne presque uniquement la rue du Rocher (elle même pourtant élargie et alignée). Visuellement, l'espace de 1775 à 1808 n'a plus rien de commun avec celui de 1836-2010.

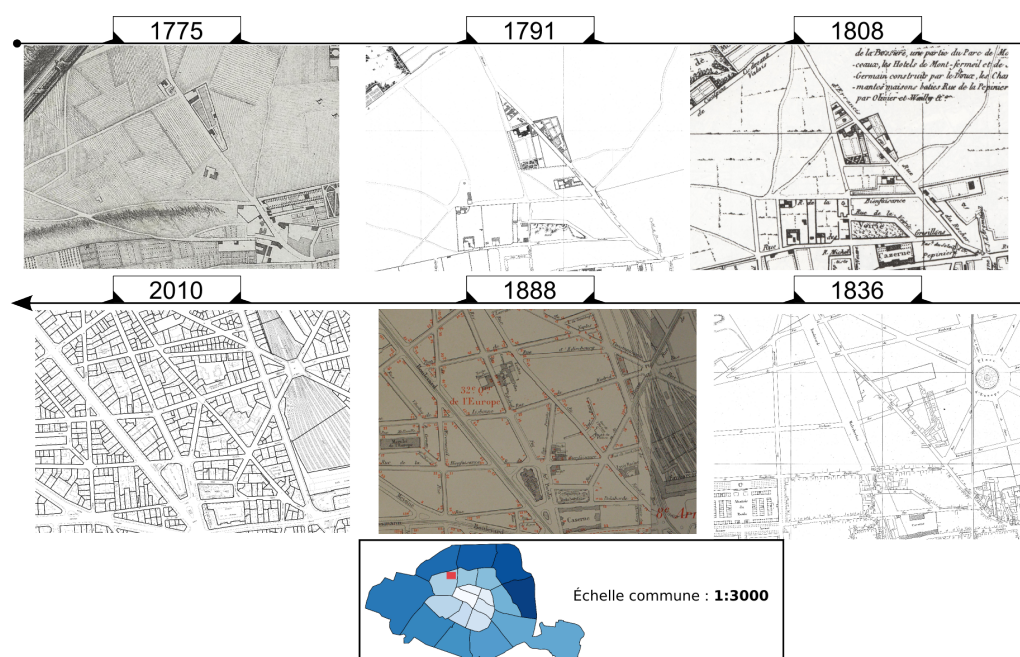


FIGURE 3.9 – Transformations entre 1775 et 2010 du faubourg parisien du Roule. Le bouleversement total de la zone se perçoit dans l'écart entre les plans de 1808 et 1836. L'espace de 2010 n'a plus rien de commun avec celui du XVIII^e siècle : presque aucun élément commun n'est repérable.

3.2.2 Principe détaillé du géoréférencement indirect

Il apparaît donc que le géoréférencement indirect est la méthode la plus adaptée pour recaler les plans topographiques de Paris. Nous la présentons en détail dans cette section. De cette façon, nous dégageons les contraintes posées par ce processus, en particulier lorsqu'il concerne des cartes anciennes.

Étapes du processus

Le principe que nous présentons ici n'est pas spécifique aux cartes anciennes. Il s'agit du processus de géoréférencement indirect tel qu'il est réalisé par la majorité

des SIG. Deux cas généraux peuvent se présenter : soit le système de projection de l'image source est connu, soit il ne l'est pas. Le schéma en figure 3.10 présente les étapes majeures d'un tel processus dans les deux cas. Il se situe après le placement de points d'amer dans l'image source (le plan ancien) et dans le repère géographique ou projeté cible. Le déroulement est le suivant :

1. on choisit des points d'amer sans ambiguïté, qui existent toujours aujourd'hui et qui sont facilement pointables,
2. les points d'amer définis dans le repère de la carte source sont projetés dans le système de coordonnées projetées de cette image - point (1) sur le schéma¹⁴. Si celle-ci n'en possède pas (parce qu'il est inconnu, ou parce qu'il s'agit d'une simple image), cette étape est ignorée. On se trouve alors dans le second cas de figure,
3. une transformation spatiale F est estimée à partir des points d'amer de l'image source et leurs correspondants dans le système de coordonnées cible. Cette étape sera détaillée dans le prochain paragraphe,
4. l'image source est transformée par F vers le système de coordonnées cible. Elle est alors potentiellement étirée et déformée.

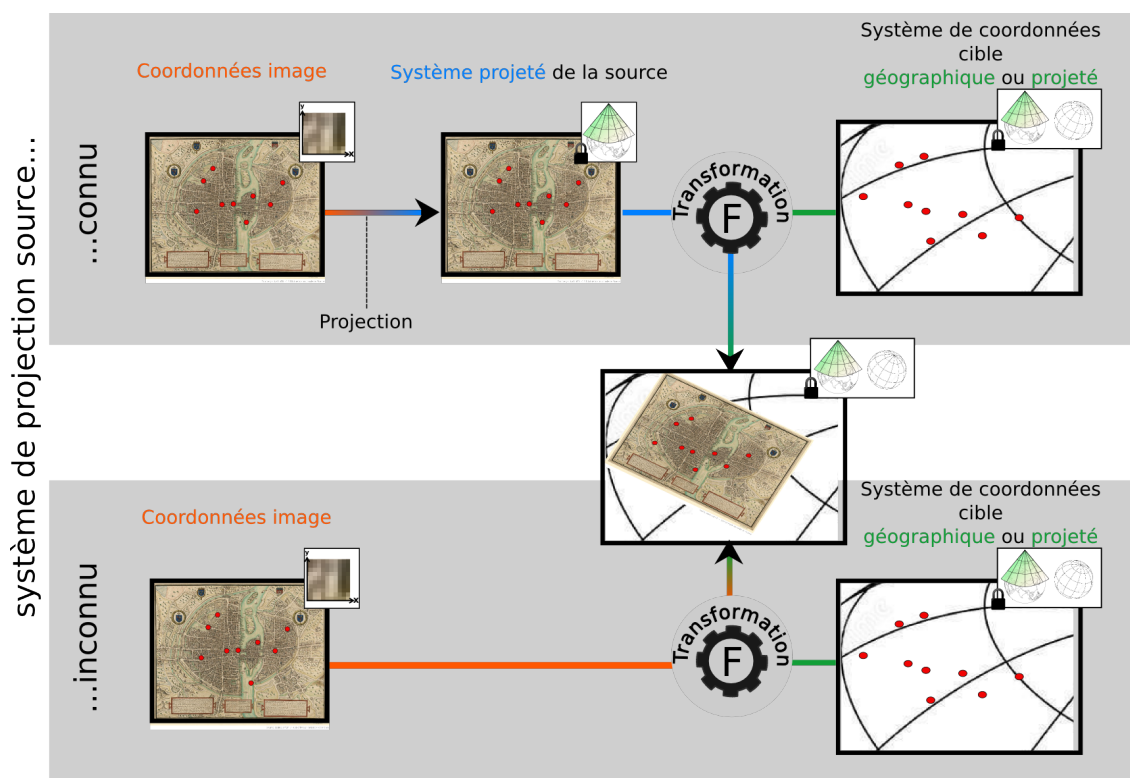


FIGURE 3.10 – Détail du processus de géoréférencement indirect

Forme de la transformation spatiale F : cas de figures possibles

14. Projeter l'image dans son système de coordonnées revient à affecter à chaque pixel sa coordonnée dans ce système. Si la carte n'est pas déformée, il s'agit d'une simple transformation affine.

La transformation de l'image source vers le référentiel cible nécessite trois étapes. D'abord, une première projection est effectuée du repère de l'image dans lequel les points d'amer sont placés vers le système de projection de la source. La transformation à appliquer à l'image est ensuite calculée en transformant ces points d'amer projetés vers leur équivalent dans le référentiel. On applique finalement cette transformation à l'image pour obtenir sa version géoréférencée.

Le procédé complet corrige donc deux types d'écart existant entre l'image source et son référentiel, qui sont décrits dans la figure 3.11. Notons que dans le cas où les systèmes de coordonnées source et cible sont connus, il est possible traiter les deux types d'écart séparément (voir le cas "connu", figure 3.10). Si les deux systèmes de coordonnées sont en outre identiques, alors seule la correction des distorsions doit être réalisée. Ainsi, les distorsions obtenues en sortie du géoréférencement sont caractéristiques des erreurs propres à la carte (erreurs de levé topographique, déformation du papier, etc.).

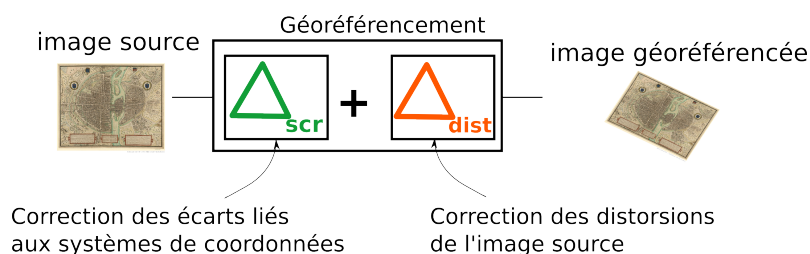


FIGURE 3.11 – Corrections géométriques effectuées lors d'un géoréférencement indirect.

Selon les connaissances sur l'image que l'on veut géoréférencer, trois cas peuvent se présenter :

1. Si F est connue *a priori*, sa modélisation est possible et applicable directement. Les points d'amer ne sont alors pas nécessaires. Il s'agit cependant d'un cas théorique, les paramètres entrant en jeu lors de la création d'une image à recaler étant rarement tous maîtrisés et jamais avec une précision absolue.
2. Si cette transformation est connue sans qu'il soit possible de la modéliser directement car incomplètement ou imprécisément définie, l'objectif est alors d'en construire une approximation que l'on ajuste à l'aide d'un ensemble de points d'amer. Cette méthode est utilisée notamment pour le géoréférencement direct de photographies aériennes, pour lequel des points peuvent être utilisés pour améliorer la précision du modèle construit à partir des paramètres d'acquisition (cf. le paragraphe 3.2.1).
3. Dans le cas où elle est totalement inconnue, on est contraint de choisir arbitrairement un modèle de transformation qu'il faut ensuite adapter au cas courant.

Dans la suite du propos, nous nous concentrerons sur le troisième cas. Pour les cartes anciennes, le modèle de transformation est toujours inconnu. En effet, même si l'on essaye de recomposer le système de projection de la carte ancienne, aucun modèle caractéristique des distorsions de celle-ci n'est disponible *a priori*.

Estimation de F dans le cas d'un modèle de transformation inconnu

Pour plus de simplicité, nous nous abstrayons momentanément des systèmes de coordonnées. Considérons simplement deux images, l'une source I_s et l'autre cible I_c sur laquelle un ensemble $\{(p_1, p'_1), (p_2, p'_2), \dots, (p_n, p'_n)\}$, $n \in \mathbb{N}$ de points d'amer sont définis, où p_i est dans le repère de I_s et p'_i dans le repère de I_c . On cherche alors une fonction de transformation F quelconque de paramètres $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$, $k \in \mathbb{N}$ telle que $F(p_i) = p'_i$.

La détermination de ces paramètres est alors réalisée par estimation statistique. Il s'agit de rechercher les paramètres estimés $\{\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_k\}$ qui définissent une fonction de transformation \hat{F} telle que $\hat{F}(p_i) = \hat{p}'_i \approx p'_i$. Cette estimation est usuellement faite par la méthode des moindres carrés ordinaires. Il s'agit alors de minimiser l'**erreur d'estimation** ou **erreur quadratique moyenne** définie par l'équation 3.1 (Pour les sources anciennes, d_2 est la distance euclidienne entre deux points). La fonction $d_2(p_i, \hat{p}_i)$ est l'**erreur résiduelle** pour l'estimation du point p_i . La figure 3.12 illustre cette erreur d'estimation.

$$RMSE(\hat{F}|F) = \sqrt{\frac{\sum_{i=1}^n d_2(p_i, \hat{p}_i)}{n}} \quad (3.1)$$

Pour se replacer dans le cadre du recalage d'image, il s'agit de minimiser $1 - RMSE(\hat{F}|F)$. Cette fonction d'erreur est précieuse car elle renseigne sur la qualité du modèle estimé, c'est à dire ici de la fonction \hat{F} . Nous verrons qu'elle peut également servir à analyser la qualité planimétrique d'une carte ancienne.

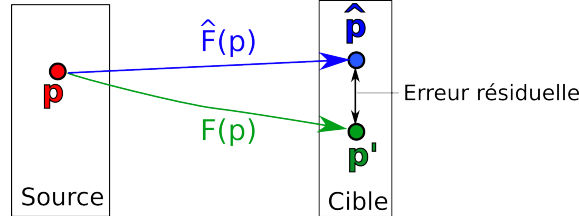


FIGURE 3.12 – Illustration de l'erreur résiduelle dans un cas d'estimation par moindres carrés ordinaires.

3.2.3 Types de fonctions de transformation dans le cas indirect

Nous sommes restés jusqu'ici dans le cadre où la fonction F n'avait pas de forme définie. L'estimation étant effectuée par moindres carrés ordinaires, les paramètres à estimer doivent suivre un modèle linéaire, ce qui limite la nature des fonctions utilisables. Nous listons ici trois modèles de fonctions de transformation utilisés habituellement dans le cadre du géoréférencement indirect (Boutoura et Livieratos, 2006). Ces méthodes sont classées par ordre croissant du nombre de points d'amer nécessaires à leur estimation. Nous présentons également succinctement une autre méthode n'utilisant pas d'estimateur statistique mais souvent utilisée pour le géoréférencement et nous jugerons de son adaptabilité aux cas des cartes anciennes. Notons enfin que récemment, des modèles d'estimation se fondant sur des noyaux gaussiens ont été utilisés avec succès pour le géoréférencement de cartes anciennes (Herrault

et al., 2013).

Dans les paragraphes suivants, on notera un point d'amer en 2 dimensions $p = (x, y)$ et nous traiterons de transformations du plan. En effet, notre système source étant déjà projeté, nous verrons dans la prochaine section comment choisir un système cible projeté adéquat aux plans de Paris. En restant continuellement dans un espace en 2 dimensions, les transformations sont plus simples et nécessitent moins de points d'amer pour être estimées. Toutes ces méthodes sont dites **globales**, c'est à dire qu'elles s'appliquent sur la totalité des points. Cependant, les transformations polynomiales sont des fonctions globales à **adaptation locale** puisque les termes des polynômes permettent de prendre en compte les distorsions locales dans un ensemble de points d'amer.

Transformation de Helmert

La transformation de Helmert est une similitude utilisée habituellement en géodésie pour effectuer des transformations entre systèmes de coordonnées géographiques. Elle est définie pour le cas général de systèmes géographiques en 3 dimension par l'équation 3.2 et comporte :

1. une translation dans chaque direction C ,
2. un changement d'échelle μ ,
3. une rotation sur chacun des axes x , y et z (R);

$$F_{helmert}(x) = C + \mu Rx \quad (3.2)$$

Dans le cas où le système source et cible sont tous deux projetés, il n'y a plus besoin que de 4 paramètres : 2 translations, une rotation et un changement d'échelle. Il faut au minimum 2 points d'amer pour estimer une telle transformation.

Transformations affines Les transformations affines sont les fonctions de transformation géométriques les plus simples. Nous présentons ici la transformation affine à 6 paramètres qui effectue deux translations, deux changements d'échelle et une rotation. La forme de cette transformation est donnée par l'équation 3.3 et nécessite au minimum 3 points d'amer.

$$\begin{aligned} F_{affine6}(x) &= xa + yb + c \\ F_{affine6}(y) &= xd + ye + f \end{aligned} \quad (3.3)$$

où a, b, c, d, e, f sont les paramètres de transformation regroupant translation (c, f) ainsi que rotation et changements d'échelle (a, b, d, e).

Transformations polynomiales

Les transformations polynomiales sont de la forme décrite par l'équation 3.4. Le nombre de points d'amer nécessaire dépend du degré m du polynôme et est donné par $\frac{(m+1)(m+2)}{2}$. Généralement, les polynomiales d'ordre 2 et 3 sont utilisées et nécessitent respectivement 6 et 10 points d'amer pour être estimées. Des polynômes d'ordre plus élevé peuvent être utilisés, mais la transformation devient alors de plus

en plus sensible aux éventuels points d'amer aberrants dus notamment à des erreurs lors du placement.

$$\begin{aligned}
 F_{poly}(x) &= \sum_{i=0}^m \sum_{j=0}^{m-1} a_{ij} x^i y^j \\
 F_{poly}(y) &= \sum_{i=0}^m \sum_{j=0}^{m-1} b_{ij} x^i y^j
 \end{aligned}
 \tag{3.4}$$

Modèle non linéaire : Thin Plate Spline

La méthode *Thin Plate Spline* est une méthode d'interpolation dans laquelle l'objectif est de construire une surface déformée de façon minimale passant par un ensemble de points. Dans le cas de points d'amer en deux dimensions, il s'agit de déterminer la fonction $F(a) = a'$ tout en déformant au minimum la carte source. Cette méthode présente l'avantage de ne pas reposer sur une estimation statistique et donc ne pas dépendre d'un nombre de points d'amer minimum (2 suffisent pour créer une surface). Cependant, elle fait correspondre exactement les points d'amer de la carte source avec ceux du repère cible et est donc particulièrement sensible aux points d'amer erronés ou imprécis. Or les cartes anciennes sont fondamentalement imprécises, et il est impossible de garantir l'exactitude des points d'amer. De plus, ceux-ci sont généralement placés de façon manuelle et donc entachés d'erreurs de pointé (Girres, 2012). Pour cette raison, nous ne considérons pas cette méthode dans la suite de ce travail.

3.3 Évaluation de la qualité d'un géoréférencement indirect

Qu'il s'agisse d'identifier des transformations par des mesures ou par la simple superposition de deux cartes, des erreurs dans cette identification peuvent apparaître lorsque les contenus des cartes sont déformés et décalés¹⁵. En plus de projeter un plan dans un espace géographique, le géoréférencement, puisqu'il transforme l'image, permet de corriger les erreurs planimétriques du plan et donc de "gommer" ses distorsions.

La qualité d'un géoréférencement se définit donc comme sa capacité à déterminer une transformation spatiale projetant les points d'amer de la source au plus près de leur correspondant dans le référentiel cible, et donc corriger les distorsions de la source. De plus, un géoréférencement dépendant d'un ensemble de points d'amer, sa qualité repose également sur celle de ces points.

3.3.1 Exactitude d'un géoréférencement

Dans le cas des cartes anciennes, les erreurs et les déformations géométriques inévitablement présentes faussent les distances entre les éléments d'un même plan comme elles affectent leur positionnement dans l'espace géographique du référentiel. Le risque principal est alors de confondre erreurs de positions et réels changements des objets représentés dans la carte ancienne (Pontius Jr et Lippitt, 2006). Lorsque

15. Les objets géographiques sont alors distordus et leur localisation erronée.

erreurs et différences de représentation se cumulent pour déformer et déplacer deux observations, le risque d'erreur d'identification et de classification d'une transition augmente d'autant plus que les objets représentés sont denses. Il est donc important que le géoréférencement corrige au mieux les erreurs planimétriques des cartes anciennes. Il s'agit donc de mesurer l'exactitude du géoréférencement, c'est à dire **le degré avec lequel il se rapproche de la "valeur vraie" - ici l'espace du référentiel.**

Cette valeur vraie n'est connue qu'au travers des points d'amer, que l'on considère fiables dans le repère cible. Cette exactitude peut se ramener aux erreurs résiduelles entre les points d'amer de la source transformés dans le repère cible \hat{a}_i par moindres carrés et les points a'_i définis dans le repère cible. On utilise généralement la RMSE(voir le paragraphe 3.2.2) pour mesurer cette exactitude.

Cette mesure présente l'avantage d'être exprimée dans les mêmes unités que les données, facilitant son interprétation. Elle est cependant particulièrement sensible aux erreurs dans les données. En effet, la mise au carré des erreurs résiduelles amplifie l'effet d'une mesure aberrante.

3.3.2 Qualité des points d'amer

Le modèle de transformation estimée par moindres carrés, et donc la précision du géoréférencement, dépend fortement des points d'amer définis. La qualité d'un ensemble de points d'amer dans la source ou dans le référentiel cible repose sur trois points : leur nombre, leur répartition et leur fiabilité.

Répartition spatiale

L'agencement des points d'amer dans l'espace de la source et de la cible est d'une importance capitale. En effet, une répartition homogène permet de réduire les erreurs induites par les déformations locales de la carte lors du calcul des paramètres de transformation. À l'inverse lorsque tous les points sont concentrés sur une zone de la carte, le modèle estimé sera juste localement mais d'autant plus faussé ailleurs que les déformations locales sont importantes. Boutoura et Livieratos (2006) illustrent parfaitement ce phénomène en comparant diverses transformations appliquées à un sous-ensemble de points définis sur une image.

Nous illustrons également ce phénomène en suivant leur méthode dans les figures 3.13 et 3.14 . Nous disposons d'une image source quadrillée que l'on souhaite projeter dans un référentiel contenant une version bruitée aléatoirement de ce quadrillage. Un ensemble de points d'amer est défini sur les intersections des lignes du quadrillage, dans la source et dans le référentiel cible. On applique ensuite un bruit aléatoire aux points du quadrillage dans le référentiel géographique 3.13. Nous appliquons d'abord une transformation rigide -ici de Helmert-utilisant tous les points de contrôle, aboutissant à la grille transformée en figure 3.14a. On applique ensuite la même transformation à l'image d'origine, cette fois en utilisant uniquement les points 15, 16, 20 et 21, menant à une seconde transformation 3.14b.

On peut ainsi constater l'importance de la répartition spatiale des points. Dans la figure a, les points répartis sur toute l'image permettent d'absorber les décalages locaux dus au bruit. A l'inverse dans la figure b seuls quatre points situés dans une

région réduite sont utilisés, aboutissant à une transformation adaptée à cette zone mais largement décalée pour les autres parties de l'image.

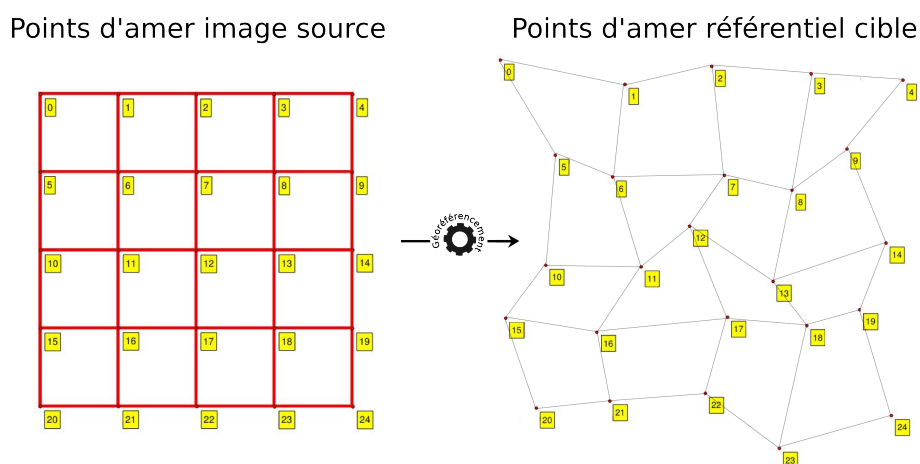


FIGURE 3.13 – Illustration des effets de la répartition des points d’amer : image source et référentiel bruité

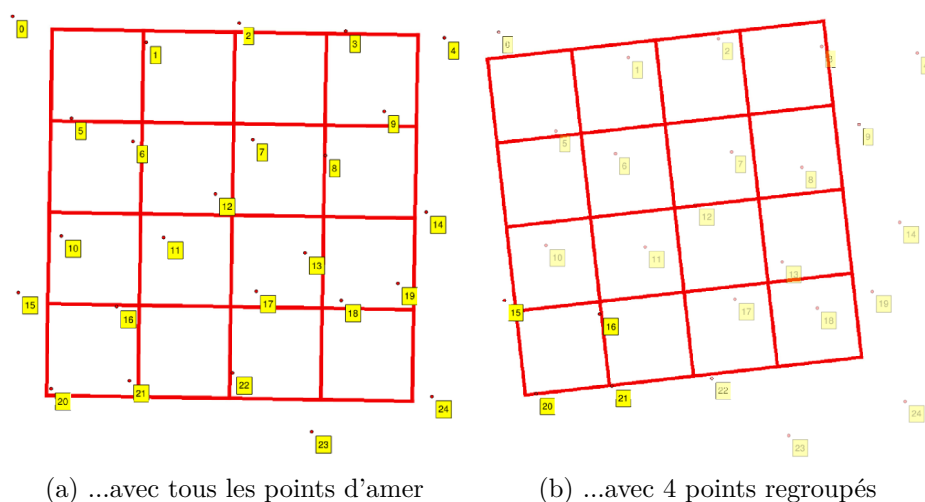


FIGURE 3.14 – Illustration des effets de la répartition des points d’amer : application

Nombre de points

En plus de leur répartition, le nombre de points placés sur la carte joue également un rôle important dans la qualité d’un géoréférencement (Boutoura et Livieratos, 2006). Tout d’abord, les différents types de transformation en imposent un nombre minimum, nécessaires pour l’estimation de leurs paramètres. Au delà de cette borne inférieure, les nouveaux points définis permettent d’améliorer l’exactitude du géoréférencement pour deux raisons :

- l’approximation du modèle de transformation est améliorée par l’ajout de nouvelles connaissances,
- la régression linéaire utilisée pour déterminer les paramètres des transformation moyenne les erreurs, ce qui permet d’absorber les déformations locales.

Cependant, il n'est pas toujours souhaitable d'augmenter ce nombre au delà d'un certain seuil. Dans la figure 3.15 issue des résultats présentés dans le travail de Herrault *et al.* (2013), on constate cependant que pour les deux distributions, le gain de précision devient nul dès 100 points placés. Une telle stagnation signifie que le géoréférencement le plus précis, étant donné la précision relative du référentiel et de la carte ancienne, a été atteint. À partir de ce moment là, l'ajout de nouveaux points devient risqué puisque qu'il peut entraîner une dégradation de la qualité du géoréférencement. Cette constatation n'est cependant vraie que pour les transformations globales sans effet local (similitudes, affines, projectives) puisque les erreurs locales moyennées se retrouvent distribuées sur l'ensemble de la carte.

Le même auteur trace également un lien entre le nombre de points et leur répartition, le choix de l'un dépendant de l'autre. Ainsi, lorsqu'un grand nombre de points est disponible, une répartition aléatoire tend à donner un meilleur géoréférencement. Inversement lorsque les repères sont faibles, une répartition régulière donne de meilleurs résultats. Ce phénomène s'explique : de nombreux points répartis aléatoirement vont mieux "couvrir" la carte qu'un faible nombre répartis de manière similaire qui laissent des zones "vides".

La règle est donc la suivante : moins on dispose de points, plus il doivent être répartis de façon homogène et régulière sur la carte.

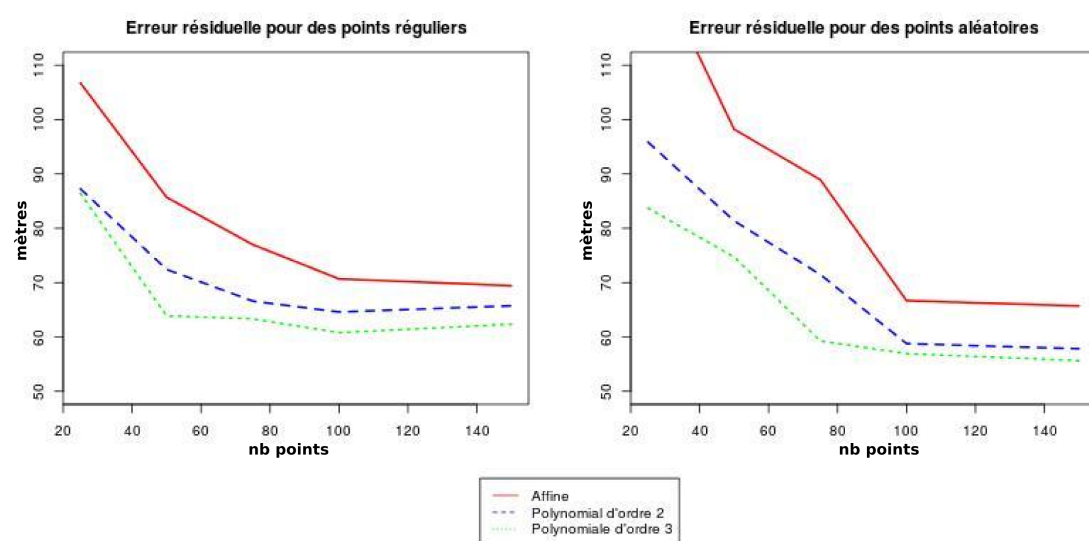


FIGURE 3.15 – Évolution de l'erreur résiduelle sur un échantillon de la "Carte de France", ici exprimée par l'Erreur Quadratique Moyenne résiduelle (RMSE) suivant différents modèles de transformation globale utilisés dans cette thèse et pour différents types de répartition de points d'amer. D'après les données issues de (Herrault *et al.*, 2013, p.41).

Fiabilité des points

À notre connaissance, la fiabilité d'un point d'appui n'a pas fait l'objet d'une définition claire parmi les différents travaux. Dans le domaine de la photogrammétrie, il s'agit généralement de recalibrer dans un repère géographique des images satellitaires ou aériennes à l'aide de relevés GPS. La fiabilité est alors directement liée à la pré-

cision planimétrique des points (Gregory *et al.*, 2005).

Puisque dans notre cas les points sont placés manuellement, on doit également considérer l'imprécision de placement ou *erreur de pointé* (Girres, 2012, p.162). Il s'agit de l'erreur accidentelle de positionnement du point résultant du processus et du choix de l'échelle de saisie : un utilisateur humain muni d'un outil de pointé (souris, etc.) va nécessairement introduire une erreur de positionnement.

Lorsqu'il s'agit de données incertaines comme le sont les cartes anciennes, il faut également s'interroger sur la confiance portée aux informations qui y sont décrites. En effet, si l'on place des points sur des objets cartographiques considérés comme "inchangés", encore faut-il être certain qu'ils le soient effectivement (Benavides et Koster, 2006). Si cette question est secondaire pour des images satellite où le référentiel (la réalité) est disponible, elle l'est moins pour des objets historiques dont la localisation peut avoir changé sans que leur forme globale n'ait été modifiée¹⁶, ou lorsque l'écart temporel entre l'objet dans la carte ancienne et celui dans le référentiel est tel qu'il devient difficile de se prononcer sur sa stabilité. Les différences de représentation entre la carte et le référentiel, les niveaux de généralisation de chacun ou encore la résolution spatiale de la carte agissent aussi sur le degré de confiance attribué à un point : plus la carte est éloignée (en termes de représentation) du référentiel, moins le point sera fiable. Zanola *et al.* (2009) a démontré cet effet sur des données 3D stéréoscopiques, le niveau de photo-réalisme d'un objet 3D influant fortement sur le degré de confiance des utilisateurs dans la qualité de ces données.

Finalement, la notion de fiabilité est dans notre cas caractérisée par deux éléments :

- la précision planimétrique des points, dépendant de la précision de la carte, de celle du référentiel, de l'échelle et du processus de saisie,
- le niveau de confiance attribué à une correspondance entre un objet de la carte et un objet du référentiel.

3.4 Proposition : une méthode de géoréférencement adaptée aux plans topographiques anciens de Paris

Dans la section 3.2, nous avons présenté les éléments nécessaires pour effectuer un géoréférencement indirect. En particulier, lorsque l'on dispose d'une image -carte, plan, etc.- à géoréférencer, cinq paramètres doivent être déterminés :

- le système de coordonnées de la source,
- le système de coordonnées cible,
- un référentiel géographique sur lequel placer des points d'amer,
- un ensemble de points d'amer dans la source et leurs correspondants dans le référentiel,
- une transformation spatiale ;

Dans cette section, nous proposons de fixer ces différents paramètres pour le cas des plans topographiques parisiens. Nous montrons notamment que l'on peut déterminer un système de projection commun pour la source et la cible, ce qui nous permet de

16. Un des meilleurs exemple est sans doute la reconstruction de bâtiments incendiés, régulièrement -du moins à Paris- modifiée par rapport à l'original malgré une certaine volonté de reproduire l'édifice "à l'identique". Citons par exemple le théâtre du Panthéon ou encore la Bourse du Commerce.

nous abstraire des erreurs de géoréférencement dues aux changements de systèmes de coordonnées et ainsi nous focaliser sur l'analyse des distorsions des plans anciens.

3.4.1 Choix du système de projection des plans topographiques anciens de Paris

En présentant le géoréférencement indirect, nous avons précisé qu'il était tout à fait envisageable de géoréférencer une image -ici une carte ancienne- sans connaissances préalables sur son système de projection. Nous avons également montré que dans ce cas précis, la fonction de transformation spatiale mêlait à la fois la projection de l'image dans le repère cible et les corrections de ses distorsions.

Il est pourtant particulièrement intéressant, lorsque l'on traite de cartes anciennes, d'être capable d'analyser leurs distorsions (Jenny et Hurni, 2011). Or il n'est possible de les isoler que si l'on connaît le système de projection de la source, ce qui permet de séparer des distorsions dues aux erreurs de la carte de celles dues à l'écart entre son système de projection et celui du référentiel cible.

Nous présentons dans cette section la méthode que nous avons adopté pour déterminer le système de projection de nos plans topographiques anciens.

Problèmes introduits par les cartes historiques

Nous avons présenté les éléments requis pour projeter un espace géographique. Nous avons dit que la reconstitution d'un système ancien est difficile, en raison surtout du manque de documentations sur la construction même de la carte. Si pour celle de Cassini le système de projection a pu être recomposé grâce aux calculs de la méridienne de Paris, retranscrits avec forts détails dans les livres édités par Cassini de Thury (voir les travaux de Berthaut (1898)), il s'agit plutôt d'une exception pour la période¹⁷. Notons toutefois que l'atlas de Verniquet est également extrêmement bien documenté puisqu'une somme importante des minutes des levés topographiques du plan ont été retrouvés et se trouvent aujourd'hui à la Bibliothèque Historique de la Ville de Paris (BHVP). La reconstitution de la triangulation de l'atlas à partir de ces minutes serait d'un aide précieuse, que nous n'avons pas traité en raison du volume de données et de leur difficulté¹⁸.

Projections des plans de Paris

La cartographie urbaine s'est dotée de systèmes locaux construits par triangulation à partir de méridiens situés à proximité. En France et à Paris, le méridien d'origine passant à l'Observatoire Royal a été utilisé dès sa détermination par Cassini en 1683. Cependant, sa position exacte a été corrigée régulièrement, la déplaçant légèrement à chaque fois. Deux calculs ont fait autorité pendant notre période :

- celui de Delambre et Méchain effectué dans les dernières années du XVIII^e siècle ;

17. On trouve le détail du système de coordonnées sur le site du projet de reconstitution et de géoréférencement de la carte de Cassini : http://cassini.ehess.fr/cassini/fr/html/5_donnees.htm. A partir du XIX^e siècle, la projection de Cassini (ou Cassini-Soldner) sera plus généralement utilisée.

18. Il s'agit, en détails, de tous les calculs et tracés des triangles et chemins polygonaux dans les rues de Paris. Pour reconstituer le canevas de l'atlas, il faut donc reproduire le travail de l'équipe d'Edmé Verniquet et replacer tous ces tracés par rapport aux triangles principaux puis à la méridienne.

- la *nouvelle méridienne de France* en 1893.

Pour Paris, tous les plans sur lesquels nous travaillons ont été construits à partir d'une triangulation fondée sur le méridien de Paris, sauf -sans certitude aucune- le plan de Maire qui s'appuie sur un travail antérieur à la détermination du méridien. De plus, Paris se trouve très exactement sur le tracé du méridien puisqu'il passe au travers de l'observatoire aujourd'hui situé dans le XIV^e arrondissement. Bien que nous ne connaissions pas en détail les méthodes de construction de chacun des plans, ceux-ci se trouvent très proches du méridien d'origine, parfaitement connu à l'époque.

Deux possibilités s'offrent à nous :

- si l'on est capable de reconstituer la triangulation du plan et que l'on connaît la position du méridien de Paris, on peut alors recréer le système de projection ancien,
- on peut essayer d'approximer le système ancien par un système projeté de référence dont l'altération linéaire sur la zone peut être considérée négligeable étant donné les imperfections inhérentes aux cartes ancienne.

Nous allons voir qu'à Paris, la seconde solution est envisageable.

Approximation du système de projection des plans de Paris par un système de projection de référence

Le choix d'un système de coordonnées adéquat dépend, on l'a vu, de la zone géographique couverte. En France, la mise en place du système géodésique de référence RGF93 (Réseau Géodésique Français) a été accompagnée par la définition d'une nouvelle projection cartographique nommée Lambert 93. S'appuyant sur l'ellipsoïde IAG-GRS80, il s'agit d'une projection conique conforme dont le méridien d'origine est Greenwich, adoptée depuis 2000 comme projection cartographique officielle pour la France métropolitaine. Précédemment à ce système existait, pour le même territoire, un ensemble de quatre projections fondées sur l'ancien système géodésique de référence NTF (Nouvelle Triangulation de la France) créée au XIX^e siècle. Les projections Lambert I, II et III couvraient ainsi la France continentale tandis que le Lambert IV avait été mis en place pour la Corse. La figure 3.16 présente les grilles de projection des quatre zones Lambert NTF et du Lambert 93 (Service de Géodésie et de Nivellement, 2009). L'objectif est donc de choisir parmi ces différents systèmes de projection celui qui présente le moins d'altérations linéaires¹⁹ au niveau de Paris. La figure 3.17 montre l'évolution de ces altérations selon la latitude. On constate que les projections Lambert I et Lambert 93 présentent des altérations faibles au niveau de Paris, inférieures au mètre. Afin de vérifier ces altérations, nous avons calculé l'altération linéaire de la borne géodésique située sur le pilier de la méridienne de France²⁰ en projetant les coordonnées géographiques (RGF93) de la borne dans les systèmes Lambert. Ces mesures ont été effectuées avec le logiciel *Circé France* de l'IGN²¹. Pour le Lambert I, nous avons obtenu une altération d'environ 6 centimètres par kilomètres²². Pour le Lambert 93, nous avons mesuré une altération

19. L'altération linéaire est l'écart existant entre par une longueur mesurée en projection et la même longueur mesurée sur l'ellipsoïde de référence.

20. Fiche géodésique IGN 75056A, point A. <http://geodesie.ign.fr/fiches>

21. <http://geodesie.ign.fr/index.php?page=circe>

22. La longueur de Paris est aujourd'hui de 18 kilomètres de l'est à l'ouest en comptant les bois de

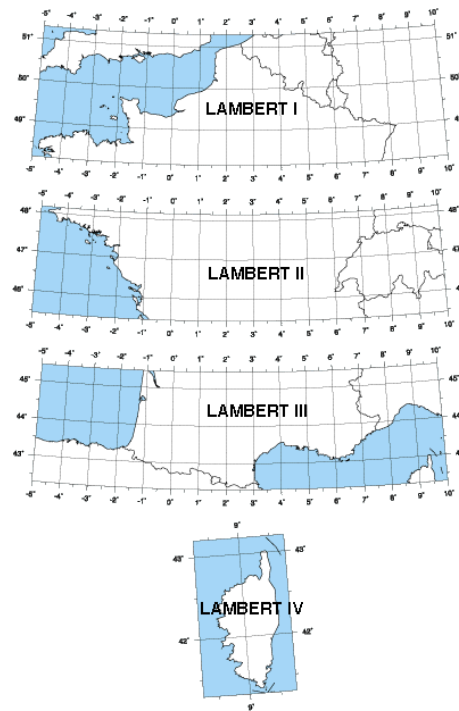


FIGURE 3.16 – Zones Lambert I, II, III et IV. Figure issue de documents IGN.

d'environ 12 centimètres.

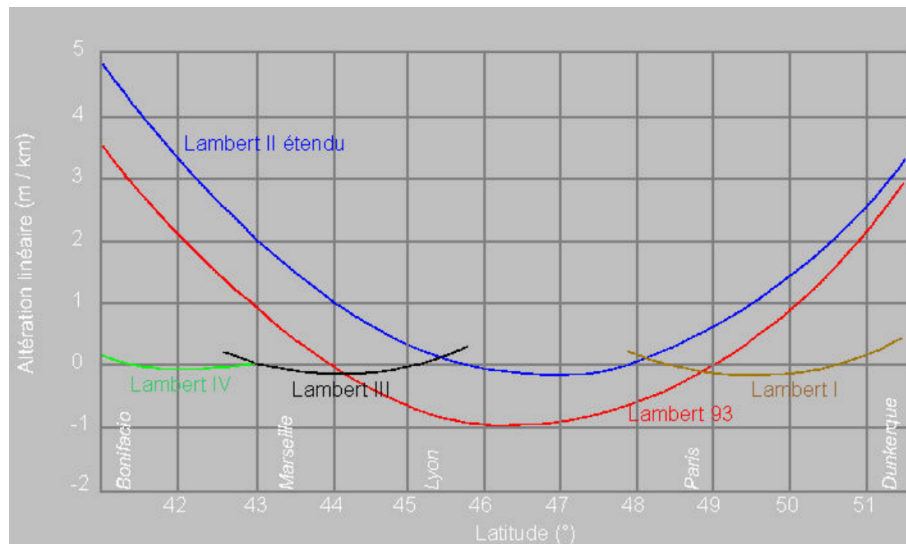


FIGURE 3.17 – Comparatif des altérations linéaires des lambert I, II, III, IV et Lambert 93. La latitude de Paris est d'environ 48,8 degrés décimaux. Figure issue de documents IGN.

Boulogne et Vincennes, et de 9,5 kilomètres du nord au sud.

Au delà de l'altération linéaire, le méridien d'origine est également à prendre en compte. En effet, les deux projections sont directes, c'est à dire qu'elles sont orientées nord-sud suivant un méridien d'origine. Nos plans présentent presque tous des carroyages alignés sur le méridien de Paris, il est donc plus simple de considérer le cas où le méridien de Paris se trouve vertical. Pour ces deux raisons, nous utilisons la projection Lambert I comme approximation des systèmes de coordonnées projeté de nos plans topographiques. Ajoutons que depuis 2006, neuf nouvelles projections coniques conformes ont été définies sur le territoire (nommées Lambert CC 42 à 50) réduisant l'altération linéaire à l'intervalle $[-9cm; +7cm]$ afin de palier les altérations relativement importantes de la projection Lambert 93. À Paris, il s'agit du Lambert CC49, qui présente une altération linéaire de 8 centimètres (Service de Géodésie et de Nivellement, 2013).

3.4.2 Choix du système de projection du référentiel géographique cible

Le choix d'un système de coordonnées pour le référentiel géographique est nécessaire pour effectuer un géoréférencement. En choisissant le Lambert I NTF comme système de coordonnées projetées des plans topographiques de Paris, nous avons déjà la possibilité, lors du géoréférencement, de séparer les écarts dus aux différences entre les systèmes de coordonnées source et cible dès lors que ce dernier est un système de coordonnées de référence. En effet, les fonctions de passage entre le Lambert I NTF et ce système sont connues.

Plus encore, le fait d'avoir considéré que le système de projection des plans de Paris pouvaient être approximé par le Lambert I NTF nous permet de choisir le même système projeté pour le référentiel cible. Ainsi, sous l'hypothèse que le Lambert I est une bonne approximation du système de projection des plans de Paris, nous éliminons les écarts dus à un changement de système lors du géoréférencement (voir la section 3.5.4). De cette façon, nous pouvons nous concentrer uniquement sur l'analyse des distorsions des plans anciens de Paris.

Sur nos plans de Paris, l'échelle est toujours indiquée dans le cadre ou le cartouche du plan, ce qui permet de transformer simplement les points d'amer placés dans l'image dans le système Lambert I. Il s'agit alors d'une simple mise à l'échelle.

3.4.3 Référentiels géographiques actuels utilisables pour le placement des points d'amer

Nous avons donc choisi un système de projection commun source-cible adapté aux plans parisiens. L'opération implique le choix d'un référentiel géographique, une représentation d'un espace commun aux données anciennes dont on connaît, pour chaque élément qui la compose, la position à la surface de la Terre avec précision (Tucci *et al.*, 2010). Généralement, des données géographiques récentes jouent ce rôle : photographies aériennes (Brovelli et Minghini, 2012; Guerra, 2000), bases de données provenant de producteurs nationaux (Plumejeaud *et al.*, 2012) ou de projets collaboratifs (Prà et Mastronunzio, 2013).

Pour obtenir un référentiel géographique utilisable pour les géoréférencer il est encore nécessaire de choisir un ensemble de données qui serviront à placer des points d'amer. À Paris, de nombreux supports peuvent fournir de telles données sous des

formes diverses : images aériennes, bases de données vecteurs ou encore cartes. Il est bien sûr possible d'utiliser plusieurs d'entre eux simultanément à la condition que tous les plans soient recalés en utilisant les mêmes données de référence de façon à conserver une précision des données homogène.

Caractéristiques requises

Plusieurs éléments permettent de choisir un ensemble de données géographiques de référence. Le critère principal allie certainement précision planimétrique des données et fiabilité de l'information représentée, qui dépendent surtout du statut du producteur de données. Ainsi en France, l'IGN, producteur officiel des bases de données de référence à l'échelle nationale doit respecter un degré de qualité et de précision défini par son statut. Ensuite, il est nécessaire de sélectionner des données représentant les mêmes objets que les cartes en suivant des règles de représentations proches. En effet, le placement précis des points d'amer n'est possible que si les deux objets ne présentent pas de différences de généralisation et de symbolisation trop importantes. Cela induit également la maîtrise des données utilisées et en particulier des spécifications de leurs saisies. Celles-ci forment des métadonnées qui doivent être diffusées en même temps que les données elles-mêmes. Parmi les classes d'objets géographiques communes à tous nos atlas parisiens, les bâtiments et les rues sont les deux plus significatives. De plus, nous avons évoqué dans le premier chapitre comment la trace de bâtiments disparus pouvaient perdurer au travers des découpages parcellaires. Il est donc important que le référentiel choisi représente ces trois éléments. Dans ce but, nous allons donc présenter les bases de données géographiques les plus pertinentes, dont des extraits sont visibles dans la figure 3.18 et en choisir une qui respecte le mieux les points abordés.

Plan cadastral français

La Direction Générale des Finances Publiques (DGFIP) assure, dans le cadre de ses missions juridiques et techniques, la mise à jour du plan cadastral couvrant le territoire national afin d'identifier, de localiser et de représenter les propriétés foncières. Sous forme papier jusqu'en 2002, le cadastre a été presque entièrement numérisé puis vectorisé pour former le Plan Cadastral Informatisé (PCI Vecteur). Les feuilles uniquement scannées forment quant à elles le PCI Image. Le cadastre vecteur a conservé la forme des feuilles du plan ancien. Il existe donc une base vecteur par commune, divisée en sections, elles-mêmes formées à partir de plusieurs sous-sections nommées *feuilles de plan*, soit au total 597 179 feuilles. Chacune de ces sections est délimitée par des objets géographiques relativement stables : cours d'eau, routes, voies de chemin de fer, etc. Le cadastre français dessine l'unité fiscale de base : la parcelle cadastrale. Parmi les plans cadastraux des communes, on distingue les plans *mis à jour*, réalisés à partir des plans napoléoniens, dont l'échelle varie entre le 1/1250^e et le 1/5000^e, et les plans dits *réguliers* construits sur de nouvelles bases²³. Puisque le cadastre a été confectionné à partir d'anciens plans, il a dû être lui-même référencé à partir des cartes au 1/25000^e de l'IGN ainsi qu'à l'aide d'orthophotographies. Afin d'assurer, malgré ces particularités, une précision

23. Les métadonnées du plan cadastral ont été diffusées par le portail GéoBretagne. Voir <http://cms.geobretagne.fr/content/fiche-de-métadonnées-type-pour-le-plan-cadastral>

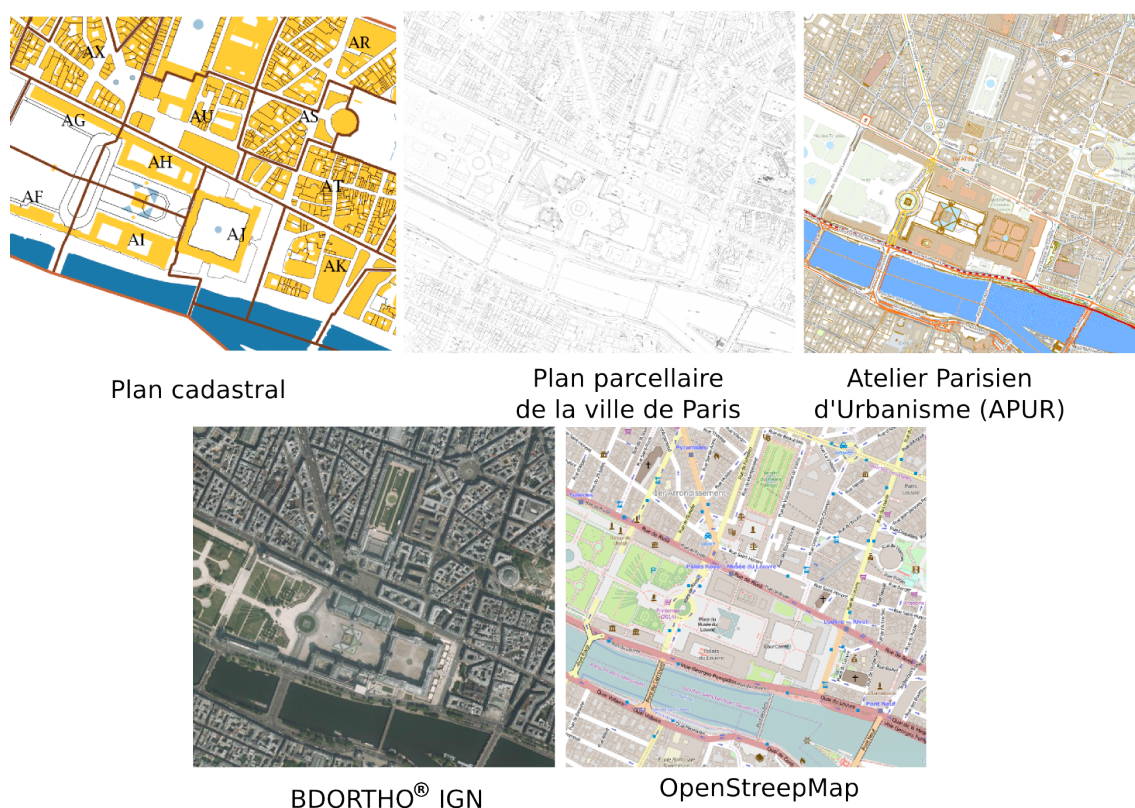


FIGURE 3.18 – Référentiels géographiques envisagés.

planimétrique suffisante pour être utilisable dans un cadre professionnel (chantiers de constructions par exemple), chaque feuille de plan s'est vu assigner une *classe de précision planimétrique*. Elle correspond à l'écart en centimètres entre les coordonnées transformées par géoréférencement indirect des feuilles et celui observé en les recalant avec une transformation affine associant une rotation et une translation. Chaque mesure de précision est ensuite adaptée d'après la précision de la base de données géographique de référence utilisée pour le géoréférencement. Ainsi, la classe de précision finale retranscrit l'exactitude de la feuille par rapport au référentiel IGN (cartes et orthophotographie) et non plus uniquement un écart interne. Le cadastre de Paris, au 1/500^e, a ainsi une classe de précision de 20 centimètres²⁴. Il n'existe cependant pas encore de métadonnées détaillées pour chaque feuille cadastrale.

Plan parcellaire de la ville de Paris

Le plan parcellaire de Paris, maintenu par la Direction de l'Urbanisme de la ville de Paris, couvre l'intégralité de la capitale à l'échelle du 1/500^e. Outre les parcelles, il représente également les limites administratives, les bâtiments et cours, les aménagements des espaces non bâtis (trottoirs, voies, escaliers, etc.) les adresses et les noms de rues et d'équipements. Le plan, diffusé en raster, vise une description extrêmement précise des structures urbaines ; il distingue par exemple les murs mitoyens des murs simples ou accolés pour chaque bâtiment. Tout comme le cadastre français,

24. Pour plus de détails sur la détermination des classes de précision, voir <http://bofip.impots.gouv.fr/bofip/5195-PGP>

il s'agit d'un plan numérisé à partir des anciennes minutes de la Ville et désormais mis à jour seulement dans sa version numérique. La précision planimétrique du plan varie entre 10 et 20 centimètres, mais peut être localement plus importante selon l'époque du levé topographique (Ville de Paris, 2008, p. 7).

Bases de donnée l'APUR

L'Atelier Parisien de l'Urbanisme est une agence regroupant divers acteurs publics d'Île de France -collectivités locales, établissements publics, État français- dont l'objectif est d'aider à la mise en place des politiques publiques urbaines par la production de divers documents d'urbanisme. En particulier, cette agence produit des bases de données géographiques sur la ville, notamment sur les parcelles cadastrales, les bâtiments et les voies de circulation. Les données créées par l'APUR sont issues des divers acteurs qui la composent, notamment, pour les parcelles et les voies, la mairie de Paris et la DGFIP, mais ne sont pas pour autant diffusées librement ou gratuitement. En particulier, les parcelles cadastrales ne sont utilisables que par les établissements associés à l'agence d'urbanisme. La précision des données APUR n'est cependant pas disponible bien que chaque couche diffusée soit accompagnée d'une série de métadonnées décrivant notamment ses spécifications de saisie²⁵.

Bases de données IGN

L'IGN propose, au sein du RGE®(Référentiel à Grande Échelle), un ensemble de bases de données décrivant le territoire français. Parmi celles-ci, la BDParcellaire® est une version du cadastre français mise en cohérence avec les autres données produites par l'établissement. Ainsi, les géométries sont corrigées et toutes les feuilles cadastrales sont rassemblées en une base de données unique. La BDParcellaire se présente sous une forme raster (c'est alors une version assemblée du PCI Image) et sous forme vecteur (créée à partir du PCI Vecteur). Outre le fait de ne plus morceler le cadastre, ces bases de données présentent l'avantage certain de disposer d'une somme de métadonnées conséquente renseignant tout le protocole de saisie ainsi que les informations de précision des données. Puisque les géométries sont recalées par rapport aux autres données du RGE, la précision planimétrique de la BDParcellaire est moindre que celle du cadastre et varie entre 0,5 mètres et 5 mètres. À Paris, elle est de l'ordre de 0.5 mètres. Enfin, elle contient, en plus des parcelles, une couche donnant le détail de toutes les surfaces bâties.

OpenStreetMap

L'initiative Open-Source OpenStreetMap vise à créer, mettre à jour et diffuser librement une base de données géographique vecteur sur la planète entière. La constitution de cette base est entièrement fondée sur un modèle collaboratif, chacun (établissement public ou privé, particulier) pouvant contribuer de façon libre. La réutilisation des données est libre, mais diffusées sous la licence contaminante ODbL, obligeant donc toute donnée dérivée à adopter la même licence. Depuis 2009, OpenStreetMap intègre une version du PCI vecteur, mettant donc à disposition les géométries de tout le parcellaire français. Contrairement aux producteurs de données officiels, le mode de mise à jour d'OpenStreetMap n'assure pas une couverture

25. Accessibles sur le Géoportail INSPIRE : <http://inspire-geoportal.ec.europa.eu/discovery/>

homogène du territoire. Le statut de capitale de Paris lui assure cependant une description très détaillée et une fréquence élevée de mises à jour. De plus, pour assouplir la saisie des objets par le grand public, un modèle de donnée très lâche a été adopté, sans aucune spécification de saisie stricte. La précision des données issues des saisies collaboratives est donc particulièrement instable.

BDORTHO® IGN

Dernier référentiel envisageable, les photographies aériennes comptent parmi les supports réguliers du géoréférencement de cartes historiques. Sur Paris, l'IGN met à disposition de façon libre, sous forme de flux WMS²⁶ ou d'images géoréférencées, la mosaïque complète de la France à une résolution de 50 centimètres. La précision planimétrique quant à elle dépend de la précision du modèle numérique de terrain sur lequel sont géoréférencées les images. Mais, sur la zone de Paris, elle reste constante à 1 mètre plus ou moins 50 centimètres (IGN, 2011, p. 9).

Sélection des données de référence.

L'utilisation du plan cadastral n'est pas envisageable pour plusieurs raisons. Si son usage est libre pour tout acteur privé ou public, individuel ou non, sa rediffusion est restrictive (elle est interdite pour les particuliers). En raison de la nature fiscale du cadastre dans lequel la mesure d'aire est cruciale, les parcelles sont aplanies pour conserver leur aire réelle malgré la projection plane du plan. Le relief est totalement absent, ce qui entraîne des décalages qui peuvent être importants dans les zones particulièrement accidentées. La BDParcellaire® présente l'avantage d'être plus fournie en métadonnées que le cadastre, et, surtout, d'être une version fusionnée de celui-ci au prix, cependant, d'une perte de précision. Une fusion de la BDParcellaire® et du cadastre en une base unique est prévue mais n'est pas, actuellement, mise en place. Cette base de données est cependant moins détaillée que les plans parcellaires de Paris et de l'APUR. Ces deux plans constituent les sources de données les plus détaillées dont on dispose sur Paris. Si le plan parcellaire de la ville de Paris est en partie mis à jour, en particulier pour les parcelles, à partir des données issues du dépouillement des feuilles cadastrales par l'APUR, lui seul est disponible librement et gratuitement. En outre, son niveau de détail et les objets géographiques représentés correspondent parfaitement à nos demandes : les bâtiments, parcelles et rues sont présentes ; le plan est à très grande échelle ; les spécifications de saisie sont connues et disponibles²⁷ ; sa précision théorique globale est de 10 à 20 centimètres. Enfin, nous n'avons pas retenu l'utilisation des photographies aériennes pour deux raisons. D'une part, la nature même des prises de vues aériennes ne permet pas d'obtenir, partout, une prise de vue à la verticale parfaite du sol. Malgré une étape de correction géométrique systématique, ces images obliques sont difficiles à comparer avec des plans, eux, parfaitement verticaux et, de plus, génèrent des zones d'occlusion (arrière des immeubles élevés, reliefs abrupts, etc.). Les parcelles y sont de plus absentes, réduisant les possibilités de placement de points d'amer. Enfin la présence du mobilier urbain, de véhicules ou de couverture végétale gêne également

26. Web Map Service, protocole permettant d'afficher dans un SIG des cartes géoréférencées provenant de serveurs de données distants. Ici, il s'agit du Géoportail <http://www.geoportail.gouv.fr/>

27. Voir <http://paris-a-la-carte-version-pl.paris.fr>

l'identification de ces points. **Nous avons donc choisi de nous appuyer sur le plan parcellaire de la ville de Paris.** En effet, son niveau de détail, son échelle et sa précision en font le meilleur candidat pour géoréférencer des plans à l'échelle de la ville. Il représente de plus tous les objets communs aux atlas parisiens, ce qui permet de maximiser le nombre d'objets pouvant servir de points d'amer.

3.4.4 Choix des points d'amer et des modèles de transformation spatiale

Types d'objets géographiques pertinents

Les points d'amer doivent être placés sur des objets géographiques communs au référentiel et à la carte ancienne. Deux catégories d'éléments sont considérés ici :

- les objets géographiques présents dans la représentation de la carte ancienne et sa représentation actuelle,
- les éléments géodésiques liés à la construction du plan tels que son carroyage, la méridienne de France et sa perpendiculaire.

Afin de permettre la comparaison des distorsions de chacun des plans anciens, il est utile de regrouper les points d'amer selon la nature des objets géographiques sur lesquels ils sont placés (Grosso, 2010). Ainsi, l'interprétation des erreurs résiduelles est plus aisée. De plus, cela peut permettre d'analyser les différents thèmes cartographiques qui composent la carte.

Les plans dont nous disposons sont pour la majorité (Verniquet, Jacoubet, Maire, Poubelle) très épurés. Les seuls types d'objets géographiques que l'on retrouve sur chacun d'eux sont :

- les bâtiments publics et religieux,
- les rues, cours et jardins,
- les cours d'eau.

Le carroyage : un support privilégié

Habituellement, on utilise des points d'amer placés sur des objets géographiques qui n'ont pas changé : bâtiments, structures naturelles, etc. Cela suppose d'être certain de leur stabilité temporelle, mais aussi de leur précision géométrique sur le plan, de l'exactitude du plan lui-même, etc. Ces incertitudes sont d'autant plus critiques que l'échelle du plan est petite.

Dans un premier temps, les plans topographiques de Paris s'orientent selon un axe sud-est/nord-ouest, plaçant la Cité et l'île Saint-Louis à la verticale (voir la figure 3.19a). Cette disposition, héritée du premier plan de Paris en vue verticale créé par Gomboust en 1652 (Gomboust, 1652)²⁸, est progressivement remplacée par une seconde dans laquelle l'île de la Cité est placée à l'horizontale (voir la figure 3.19b). Cette dernière laisse enfin place au XVIII^e siècle à une orientation nord-sud fixée par la méridienne de France (figure 3.19c). Cette évolution n'est toutefois pas homogène : le plan de Jaillot (1775) est par exemple encore orienté d'est en ouest, et celui en perspective de Bretez (1739) selon l'axe sud-est/nord-ouest.

Sur les plans topographiques, un carroyage est souvent dessiné pour aider à se repérer. Sur de nombreux plans, la position du carroyage et les dimensions des cellules sont arbitraires et sans rapport avec l'espace géographique représenté. Cependant,

28. Lui-même reprenant l'orientation des plans en vue cavalière qui le précèdent.

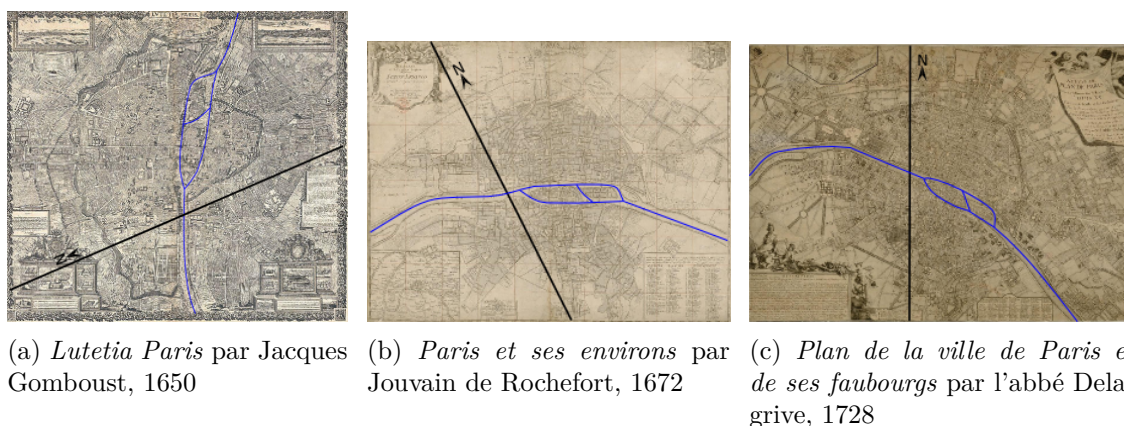


FIGURE 3.19 – Différentes orientations des plans de Paris aux XVII^e et XVIII^e siècles. La méridienne apparaît en noir. Le tracé de la Seine est figuré en bleu pour aider la lecture.

sur les atlas de Verniquet, Jacoubet et l'atlas municipal, ce carroyage s'appuie sur le tracé de la méridienne qui devient une des lignes verticales. Sur l'atlas Verniquet, la perpendiculaire de l'Observatoire sert également de ligne horizontale²⁹. Toutes les lignes verticales du carroyage sont donc des parallèles à la méridienne, et toutes les lignes horizontales des parallèles à la perpendiculaire dans la carte. Enfin, les dimensions des cellules dans l'espace géographique sont connues et notées sur la carte. En reconstruisant ce carroyage dans le référentiel géographique, nous pouvons l'utiliser pour placer des points d'amer indépendants des objets géographiques représentés et de leurs distorsions. La figure 3.20 présente le carroyage du plan en rouge aligné sur la méridienne et sa perpendiculaire apparaissant en vert et se croisent au niveau de l'Observatoire Royal. L'alignement du carroyage sur les repères géodésiques que sont la méridienne et sa perpendiculaire fait de celui-ci un canevas au même titre qu'une triangulation. Il est de plus indépendant de la représentation cartographique, des erreurs géométriques liées au levé des objets géographiques et des incertitudes sur la stabilité temporelle des points d'amer. Pour ce qui concerne les éléments géodésiques, tous les plans sauf le cadastre de Vasserot présentent un carroyage aligné sur la méridienne.

Types de points d'amer choisis

Comme nous l'avons expliqué dans la section 3.3, les points d'amer doivent être choisis de façon à être nombreux, fiables et répartis de façon homogène dans la carte ancienne. Ainsi, les cours d'eau et jardins, par leur nature instable dans le temps, ne constituent pas de supports fiables pour des points d'amer. Par exemple, les berges de la Seine ont été fortement transformées depuis la fin du XVIII^e siècle. Les rues, et en particulier leurs bords, ont également changé au fil des opérations d'alignement. Elles constituent pourtant des éléments visuellement saillants sur lesquels il est facile de placer des points d'amer. Il en est de même pour les bâtiments, qui peuvent avoir évolué (extension, reconstruction à l'identique, etc.) mais dont l'identification

29. La perpendiculaire ne correspond pas au parallèle automécoïque du Lambert I, ni à aucun parallèle géographique. Cependant, elle est une bonne approximation du parallèle passant par l'observation en raison de la faible largeur de la ville.

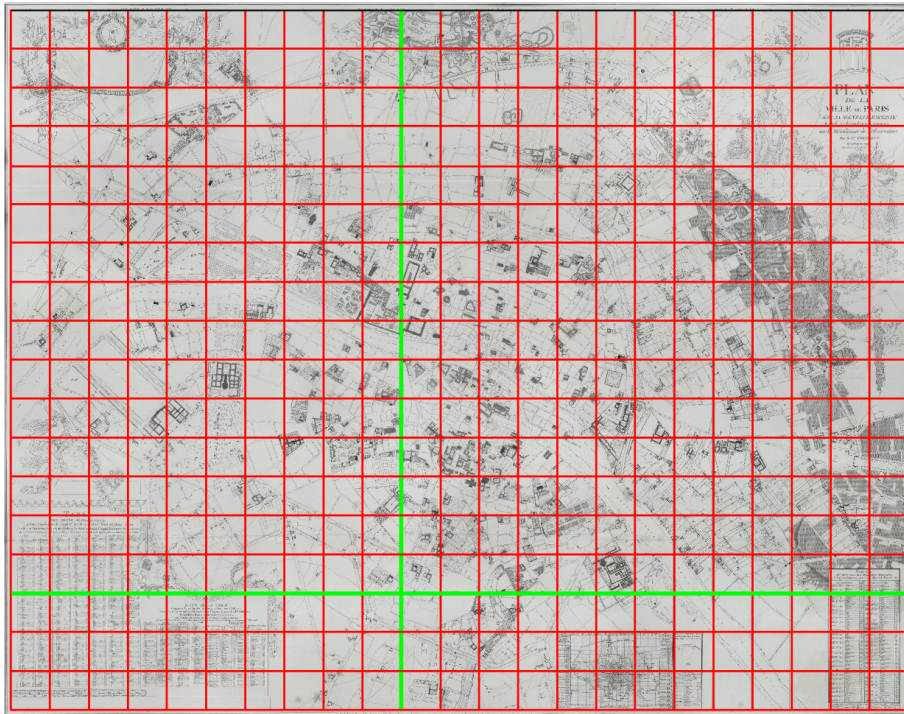


FIGURE 3.20 – Le carroyage de l’atlas de Verniquet (ici en version réduite au 1/8^{ème}) aligné sur la méridienne de Paris et sa perpendiculaire.

est aisée. Enfin, les éléments géodésiques (tracés de la méridienne, des points de triangulation, etc.) sont pour leur part jugés fiables, en particulier sur des plans dont un objet principal est l’exactitude géométrique. De plus, de par la nature du carroyage, il est possible de répartir des points d’amer de façon très homogène sur celui-ci (aux intersections des lignes du carroyage).

Nous avons conservé comme support les points d’amer les bâtiments publics et religieux, les rues et le carroyage des plans. Les points d’amer sont placés sur eux de la façon suivante :

- sur les rues, sur la bordure délimitant leur tracé aux intersections et changement de direction (voir la figure 3.21a,
- sur les bâtiments aux angles des murs extérieurs, 3.21b
- sur le carroyage aux intersections des lignes et colonnes de la grille 3.21c.

Enfin, l’union de plusieurs ensembles de points d’amer placés sur des objets géographiques de natures différentes peut permettre d’obtenir un nombre élevé de points couvrant toute la carte. Il peut être en revanche pertinent de n’utiliser qu’une seule catégorie d’objets géographiques de sorte à produire un géoréférencement fondé sur des points d’amer de fiabilité homogène. Comme l’explique (Grosso, 2010, p. 111) -pour le cas précis de la carte de Cassini-, le regroupement des points d’amer par classe d’objet géographique et l’utilisation de l’ensemble générant l’erreur résiduelle la plus faible permet d’obtenir un géoréférencement dont les erreurs résiduelles sont plus facilement interprétables.

Modèles de transformation spatiales considérés

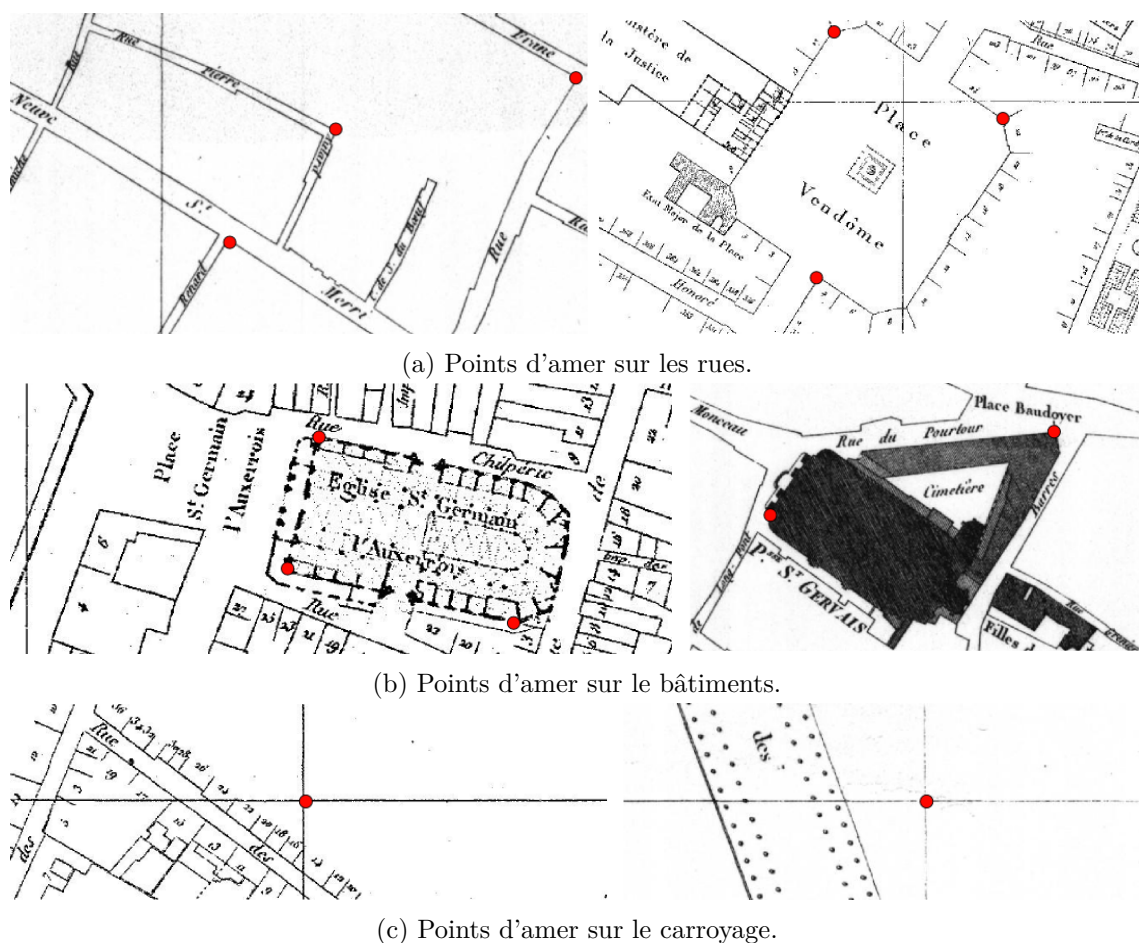


FIGURE 3.21 – Types de points d'amer considérés. Les échelles ne sont pas constantes.

En l'absence d'hypothèse sur les distorsions des plans de Paris, nous ne pouvons juger a priori de la pertinence des modèles de transformation présentés dans la sous-section 3.2.3. C'est pourquoi nous allons tester toutes ces méthodes afin d'évaluer a posteriori leur pertinence.

3.4.5 Conclusion sur l'approche de géoréférencement proposée

Pour géoréférencer les plans de Paris dont nous disposons, nous proposons donc d'aligner préalablement les systèmes de projection du plan et du référentiel géographique sur lequel doivent être placés les points d'amer. Cela nous permet alors de nous abstraire d'éventuelles distorsions liées aux écarts entre les systèmes source et cible et ainsi de nous concentrer sur les distorsions propres du plan. Celles-ci peuvent être causées par des erreurs lors des opérations de levé topographique du plan, lors de l'agencement des différents levés dans le plan final. Elles peuvent finalement être la conséquence de déformations du support même du plan en raison du vieillissement du papier et des conditions de conservation. Le schéma présenté dans la figure 3.22 illustre le processus de géoréférencement des plans topographiques de Paris que nous utilisons dans cette thèse.

Connaître les distorsions d'un plan ancien est utile, en particulier lorsque l'on veut

lier les objets géographiques qu'il représente avec d'autres plans également géoréférencés. En effet, en raison de leurs distorsions, les objets géographiques représentés sur différents plans se trouvent décalés ou déformés une fois géoréférencés, gênant alors la reconnaissance des transformations ou des persistances et la liaison des différentes représentations des entités géographiques au cours du temps. La connaissance de ces distorsions permet de les corriger ou d'aider le processus de liaison.

Nous avons également la possibilité de nous appuyer sur le carroyage du plan afin de procéder à son géoréférencement. Ce support, sur lequel peuvent être placés des points d'amer, a l'avantage de s'abstraire des distorsions liées aux levés topographiques. Il reste cependant toujours sujet aux distorsions liées à la conservation du plan. Géoréférencer un plan grâce à son carroyage nécessite cependant de recréer celui-ci dans le référentiel cible. Dans notre cas, nous avons vu qu'en nous plaçant dans le système Lambert I NTF, la construction d'un carroyage est aisée puisque son axe vertical se trouve aligné avec le méridien de Paris. Les points d'amer placés sur le carroyage peuvent donc former un ensemble particulièrement fiable si le plan lui-même s'appuie sur un canevas géodésique de qualité.

D'autres types de points d'amer possibles, placés sur des objets géographiques partagés par tous les plans que nous traitons, ont été identifiés. Le choix d'un type de points, géodésique (carroyage) ou géographique (bâtiments, rues) dépend du plan. Il est cependant préférable que la même classe de points d'amer soit utilisée pour l'ensemble des plans.

De même, le modèle de transformation spatiale doit être choisi pour chaque plan de façon à obtenir le géoréférencement le plus exact possible.

Afin de tester et valider notre approche, nous l'appliquons au plus ancien de nos plans : l'atlas de Verniquet.

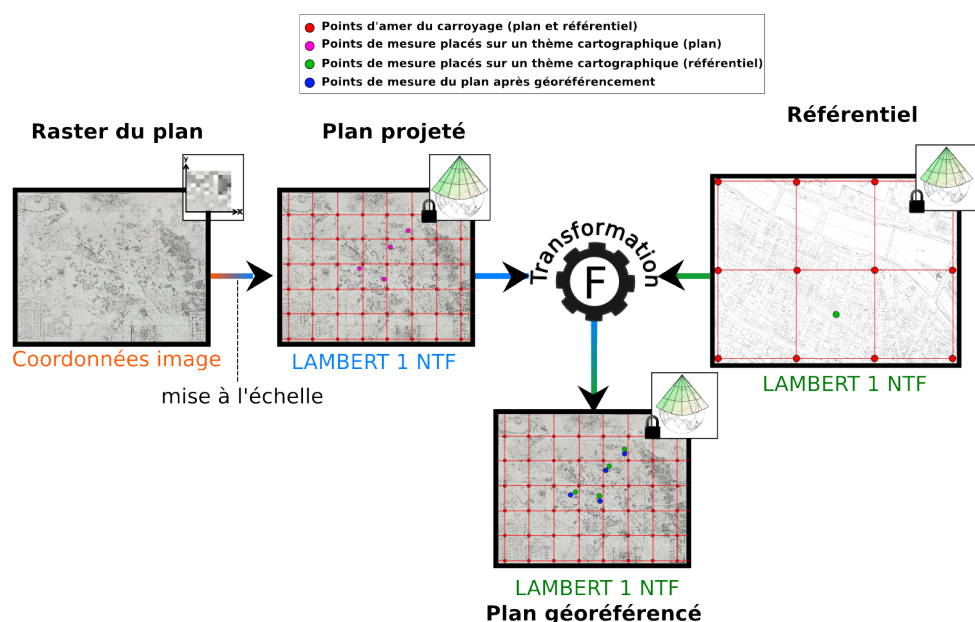


FIGURE 3.22 – Schéma du processus de géoréférencement des plans anciens de Paris. Les points d'amer sont en rouge. L'erreur résiduelle des points de mesure du plan (rose) une fois géoréférencés (bleu) permet de mesurer les erreurs d'un thème cartographique.

3.5 Cas applicatif : l'atlas de Verniquet

L'objectif de cette section est de valider l'approche de géoréférencement des plans anciens de Paris présentée dans la section précédente sur l'atlas de Verniquet. L'atlas de Verniquet est un ensemble de 72 feuilles couvrant l'intégralité de Paris dans ses limites pré-1860 ainsi que les faubourgs environnant (voir la figure 3.23). Très épuré, il contient essentiellement le tracé des rues, nommées, ainsi que les bâtiments des Biens Nationaux. Cet atlas a été de plus construit avec le souci constant de l'exactitude géométrique, et s'appuie sur un carroyage aligné verticalement avec la méridienne de France et horizontalement avec sa perpendiculaire passant à l'Observatoire Royal. Enfin, il est le dernier grand atlas de Paris présentant la ville de l'Ancien Régime, avant les grandes transformations du XIX^e siècle. En raison de ces propriétés particulières (précision, statut unique), nous le choisissons pour valider notre approche de géoréférencement.

Nous montrons tout d'abord comment il est possible de reconstruire facilement, dans le référentiel géographique, le carroyage du plan en s'aidant de quelques repères géodésiques tels que la méridienne et sa perpendiculaire. Cette construction est abordée dans la section 3.5.1.

Ensuite, dans la section 3.5.3, nous appliquons notre approche de géoréférencement et nous choisissons la classe de points d'amer ainsi que le modèle de transformation permettant d'obtenir un géoréférencement d'une grande exactitude.

Dans la section 3.5.4, nous présentons une méthodologie -et son application- permettant de vérifier l'hypothèse de l'approximation du système de projection du plan par le Lambert I NTF.

Dans une dernière section (3.5.5), nous proposons d'aller plus loin en utilisant les autres types de points d'amer pour analyser les distorsions du plan. Ceci nous permet de fournir une analyse qualitative et quantitative de la qualité de l'atlas de Verniquet. Nous explorons également la possibilité d'utiliser l'atlas de Verniquet comme référentiel géographique pour le géoréférencement de plans de Paris antérieurs au XIX^e siècle.

Nous concluons finalement sur l'approche proposée et son application aux autres plans de Paris dont nous disposons.

3.5.1 Construction du carroyage

L'objectif est donc de reconstruire le carroyage du plan dans le référentiel géographique, afin que l'on puisse y placer des points d'amer qui permettront d'estimer une transformation spatiale et géoréférencer les différentes feuilles de l'atlas de Verniquet. Nous effectuerons en section 3.5.4 des tests afin de vérifier la validité de l'utilisation du carroyage comme support des points d'amer.

Placement de la méridienne et de sa perpendiculaire dans le référentiel géographique projeté

Afin de construire ce carroyage, la première étape consiste à placer dans le référentiel la méridienne de France et sa perpendiculaire représentées dans l'atlas. Le plan comme le référentiel étant considéré comme projeté suivant le système Lambert

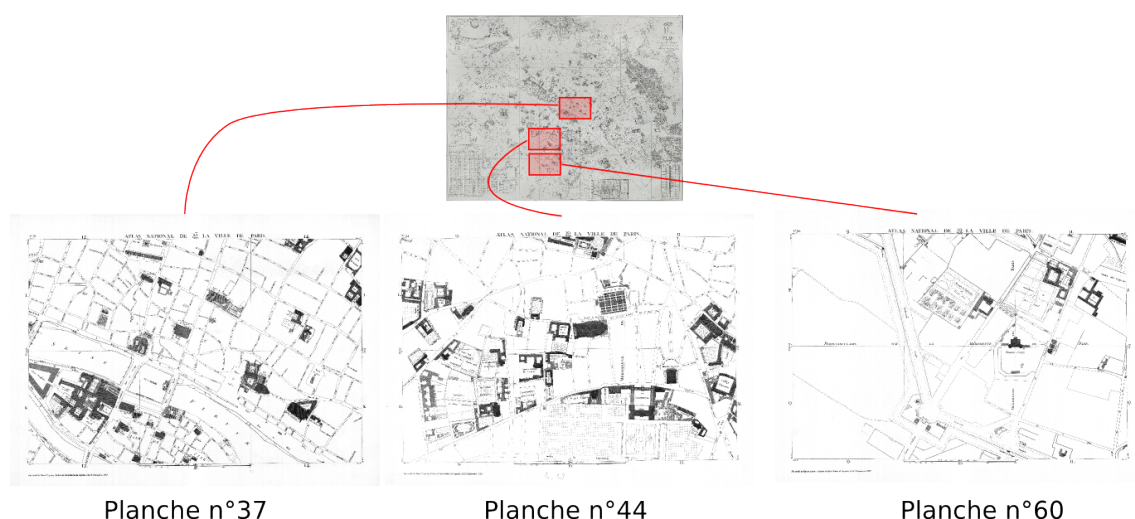


FIGURE 3.23 – Zoom sur quelques planches de l’atlas de Verniquet

I NTF, il suffit de déterminer la position de la méridienne dans celui-ci. Le méridien d’origine du système Lambert I étant le méridien de Paris, il correspond à la ligne droite verticale d’équation $x = 0$. Cependant, le Lambert I décale de 600000 mètres les coordonnées en x pour éviter les coordonnées négatives à l’ouest du méridien d’origine. Finalement, le méridien de France peut être simplement tracé par la droite $x = 600000$.

Placement de la perpendiculaire à la méridienne

La perpendiculaire à la méridienne ne correspond à aucun objet géodésique précis puisqu’il s’agit d’un repère servant uniquement à la création du carroyage. À la différence de la méridienne, sa perpendiculaire ne correspond pas à un objet géodésique de référence. Afin de placer cette dernière, nous nous sommes appuyé sur les minutes de l’atlas de Verniquet, lesquelles contiennent le plan détaillé de l’Observatoire Royal et les tracés de la méridienne et sa perpendiculaire. Ce plan détaillé, issu des travaux de Pronteau (1986), est présenté en figure 3.24. Afin de déterminer la position de cette perpendiculaire, nous nous sommes appuyé sur deux mesures. D’une part, nous avons géoréférencé le plan détaillé de l’Observatoire en nous appuyant sur le plan parcellaire de la ville de Paris, puis nous avons tracé la droite horizontale superposée au tracé de la perpendiculaire ainsi géoréférencé. D’autre part, nous avons construit cette même perpendiculaire dans le référentiel à partir des cotes indiquées sur le plan détaillé de l’Observatoire. En effet, celui-ci présente une série de mesures de distances, en toises, entre la perpendiculaire et les murs intérieurs du bâtiment. Ces mêmes mesures ont été utilisées sur le plan parcellaire de la ville de Paris pour reconstituer la perpendiculaire. L’erreur de placement final, entre ces deux perpendiculaires, est d’environ 1.2 mètres. Nous avons finalement choisi de conserver la moyenne des positions de ces deux mesures.

Construction de la grille

La construction du carroyage final est alors simple. L’atlas de Verniquet indique, sur la bordure des feuilles, la largeur des carreaux : 100 toises. La correspondance de

Observatoire des minutes de Verniquet

Observatoire BDORTHO®



FIGURE 3.24 – L’observatoire de Paris, dans les minutes de l’atlas de Verniquet et dans une orthophotographie actuelle (BDORTHO®)

cette ancienne mesure³⁰ avec le mètre est de 1,9490625 mètre pour une toise. Les carreaux de l’atlas de Verniquet mesurent donc 194,9 mètres de coté. La grille peut être ensuite créée en traçant les parallèles à la méridienne et à sa perpendiculaire avec un pas constant. Le carroyage finalement construit est visible dans la figure 3.25

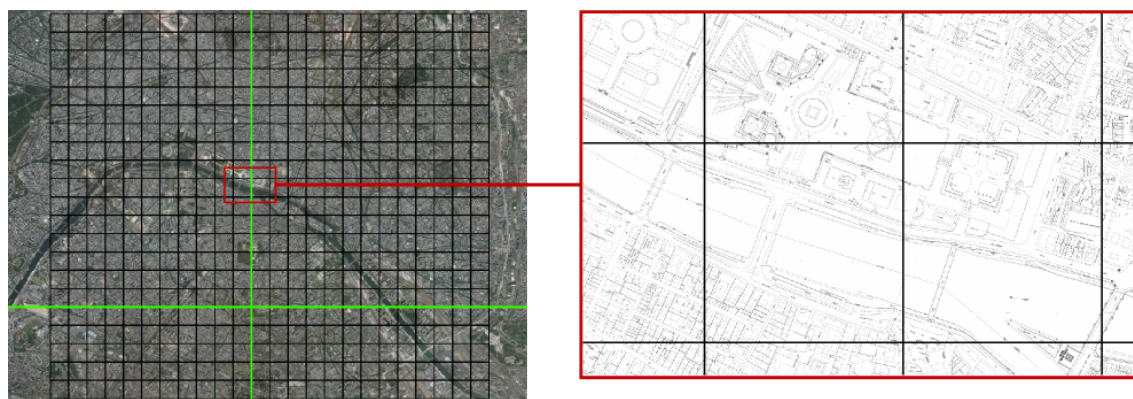


FIGURE 3.25 – Carroyage de Verniquet dans le référentiel géographique (BDORTHO à gauche, plan parcellaire de Paris à droite)

3.5.2 Placement des points d’amer sur les différents thèmes cartographiques et géodésiques

Rues

779 points d’amer ont été placés sur les rues. Puisque la fiabilité de ces points est un élément important de la qualité d’un géoréférencement, nous avons suivi quelques

30. Il s’agit ici de toises du Châtelet

règles de placement. En effet, il est important de s'assurer que ces points sont situés sur des objets n'ayant pas -ou peu- changé. Or, les rues sont des objets dont les transformations peuvent se produire sans être aisément perceptibles sur un plan, comme c'est le cas lors d'alignements. Deux règles ont été suivies :

- les points doivent être placés sur des parties visuellement saillantes de la rue, sur l'atlas comme sur le référentiel géographique, de façon à minimiser les erreurs de pointé,
- les points doivent être situés sur des rues dont la forme n'a pas -ou peu- changé jusqu'à aujourd'hui.

Pour respecter la première règle, nous avons placé les points uniquement sur les coins des rues (voir la figure 3.21a) et aux intersections, la saillance visuelle de ces deux éléments limitant les risques d'erreur. Pour nous assurer de la stabilité temporelle des points placés, nous nous sommes reporté au dictionnaire des rues fourni par la Mairie de Paris Ville de Paris (2013) et au dictionnaire historique des frères Lazare renseignant sur les transformations des rues au XIX^e siècle (Lazare et Lazare, 1844, 1855). Ces dictionnaires présentent en effet l'historique de toutes les rues de la ville et la liste de leurs transformations. Le dictionnaire de la ville de Paris ne renseigne cependant pas sur leurs élargissements.

Bâtis

Sur le bâti, 523 points d'amer ont été localisés sur les bâtiments historiques de la ville actuelle. Les points ont été placés sur les angles des murs extérieurs des bâtiments, de façon à minimiser l'erreur de pointé.

Carroyage

L'atlas étant constitué de 72 feuilles, les points d'amer du carroyage ont dû être définis pour chacune d'entre elles. Ainsi, 446 points ont été placés aux intersections des lignes et colonnes du carroyage.

3.5.3 Choix du type de points d'amer et d'une transformation spatiale

L'objectif est de déterminer la transformation et l'ensemble de points d'amer qui aboutissent au géoréférencement le plus exact, selon la définition de l'exactitude donnée dans la section 3.3. L'erreur quadratique moyenne (RMSE) est une mesure de l'exactitude du géoréférencement. Sa précision peut quand à elle être critiquée en observant la dispersion des erreurs résiduelles. L'atlas étant composé de 72 planches, les points d'amer sont placés indépendamment pour chacune d'entre elles. Généralement, l'évaluation d'un géoréférencement s'effectue pour une feuille unique. Le recalage global de l'atlas est plus intéressant. Pour ce faire, nous avons géoréférencé chacune des planches et calculé l'erreur quadratique moyenne résultante. Ces erreurs par planche permettent d'étudier la qualité globale du géoréférencement :

- leur répartition dans l'espace renseigne sur les distorsions globales de l'atlas³¹,
- la moyenne des erreurs de chaque planche donne l'exactitude globale du géoréférencement,
- les écarts entre erreurs par planche renseignent sur sa précision.

31. Déformations dues aux presses d'imprimerie, erreurs générales de levé ou dans les points d'amer, etc.

Qualité selon la transformation et l'ensemble de points d'amer

Un processus de géoréférencement repose sur un ensemble de points d'amer et une transformation géométrique. Nous avons considéré les quatre transformations suivantes communément utilisées (Jenny et Hurni, 2011; Boutoura et Livieratos, 2006) :

- transformation de Helmert à 4 paramètres,
- transformation affine à 6 paramètres,
- transformations polynomiales d'ordre 2 et 3.

La méthode *Thin Plate Splines* est écartée puisqu'elle fait correspondre exactement les points d'amer. Chaque transformation a été appliquée aux trois ensembles de points définis précédemment. Précisons également que si le plan est composé de 72 planches, plusieurs d'entre elles ne contiennent pas d'objets géographiques (tables de renseignements, cartouches, décorations) mais possèdent tout de même un carroyage. Pour ne pas fausser la comparaison entre les points placés sur le carroyage et les autres, ces planches ont été ignorées. De plus, le nombre de points d'amer minimal varie selon la transformation, rendant le géoréférencement de certaines planches impossible lorsque le nombre de points est trop faible. Enfin, lorsqu'une planche dispose du minimum de points d'amer, l'erreur résiduelle sur chaque point est nécessairement nulle, ce qui tend à diminuer l'erreur quadratique sur l'atlas entier. Dans ce dernier cas les paramètres peuvent être calculés mais il y a alors un fort risque de sur-ajustement, accentué par le manque de précision de nos points. L'erreur standard d'estimation étant alors nulle, ces planches ne permettent pas d'estimer la précision et l'exactitude d'une estimation. Nous les éliminons donc dans un premier temps, au même titre que les planches contenant trop peu de points. La figure 3.26 fait apparaître la répartition des points d'amer par planche selon le type de points et de transformation. Les planches grisées ne possèdent pas assez de points pour estimer les paramètres de transformation, et celles en orange possèdent exactement le nombre de points minimal. Naturellement, seuls les points placés sur le carroyage couvrent l'atlas de façon homogène et suffisent presque toujours à estimer les paramètres des transformations de chaque planche. Il y a 12 points par planche, sauf sur celles du bord gauche qui n'en comptent que 9, et sur lesquelles une transformation polynomiale d'ordre 3 n'est pas applicable. Les deux autres planches invalides pour cette transformation sont le résultat d'un carroyage localement illisible ou effacé.

Aucun des deux autres ensembles de points ne couvre entièrement l'atlas ni même la totalité de l'intérieur de Paris. La situation reste cependant relativement stable avec l'augmentation de la complexité des transformations, et reste envisageable pour la majorité des planches dans le cas des transformations affines. Les points placés sur les bâtiments couvrent peu l'est de la ville, ce qui s'explique d'une part par la faible densité de bâtis religieux et publics par rapport au reste de Paris, et à la plus grande transformation de la Ville qui empêche le placement de points sur un référentiel récent. Aucun des ensembles de points géographiques (rues et bâtis) n'est satisfaisant pour l'extérieur de Paris. L'utilisation du carroyage semble donc être la meilleure solution. Il faut cependant vérifier que celui-ci est fiable et permet un géoréférencement de qualité.

Mesures de la qualité du géoréférencement

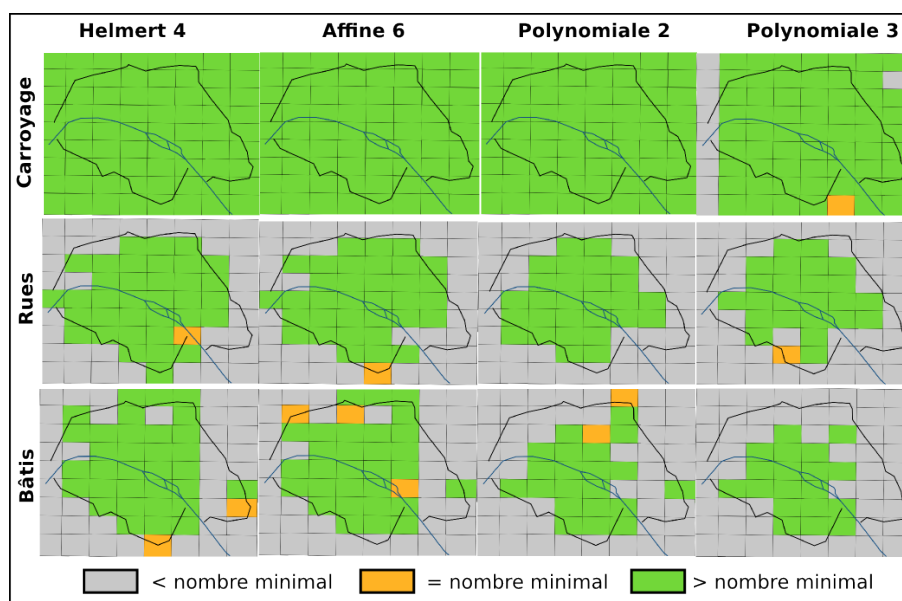


FIGURE 3.26 – Nombre de points d’amer par planches de l’atlas de Verniquet et applicabilité des transformations géométriques.

Les résultats des mesures sont visibles dans la figure 3.27, qui fait apparaître la moyenne de l’erreur quadratique moyenne pour l’ensemble de planches de l’atlas ainsi que la dispersion de ces erreurs pour chaque type de points et chaque transformation. La moyenne des RMSEs fournit un indicateur de l’exactitude du géoréférencement. Leur dispersion renseigne quant à elle sur la précision du géoréférencement, c’est à dire sur la variabilité autour des valeurs vraies. Nous délaissions donc les mesures propres à chaque planche pour s’occuper de l’ensemble de l’atlas. Ceci présuppose l’absence de points d’amer aberrants au sein des planches, qui peuvent biaiser localement l’estimation des transformations. Afin de nous assurer de cela, nous avons appliqué un test de Grubbs (Grubbs, 1950) en parallèle du géoréférencement de chaque planche, qui consiste à vérifier si la distribution des résidus suit une loi normale, hypothèse de base de la méthode des moindres carrés. Sur les trois ensembles de points considérés, aucun point ne s’est révélé être une mesure aberrante. La figure 3.27 fait d’abord apparaître la faible amplitude des erreurs résiduelles. **Ainsi, même dans le pire des cas, la RMSE moyenne sur l’atlas n’est que de 3,2 mètres, et le plus fort écart-type de 1 mètre. Quelque soit l’ensemble de points d’amer choisi, l’atlas de Verniquet peut donc déjà être considéré comme un plan de grande qualité.** La même figure montre également que les erreurs quadratiques moyennes sont nettement inférieures pour les points placés sur les rues et les bâtiments, ne dépassant pas 1 mètre. Dans le cas de l’ensemble *Bâtis*, leur dispersion autour de la moyenne augmente tandis que les transformations prennent en compte les distorsions locales. Cet effet s’observe moins pour le cas des rues, et est même inverse pour le carroyage (à l’exception de la polynomiale d’ordre 3). Pour le cas du bâti, l’exactitude diminue tandis que la précision stagne. Ceci est un effet de la répartition des points d’amer en petits groupes distants les uns des autres. Les planches contiennent souvent très peu de

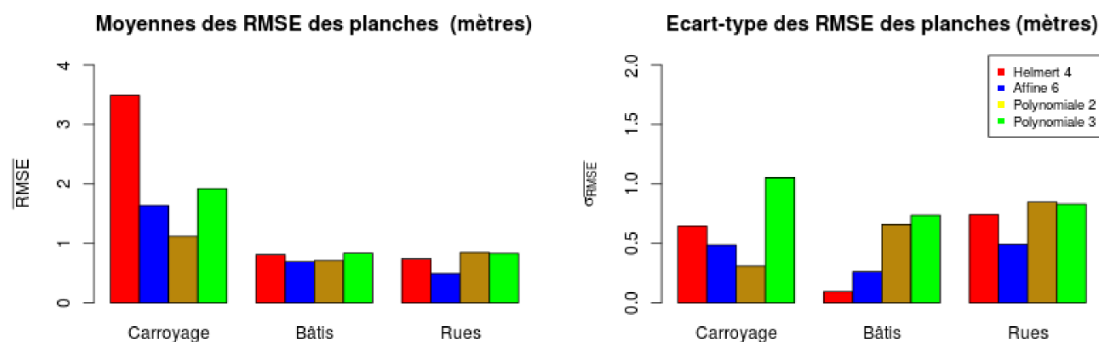


FIGURE 3.27 – Moyenne et écarts-types des erreurs quadratiques moyennes sur les Planches de l'atlas de Verniquet.

bâtiments -parfois un seul-. Sous l'hypothèse que les distorsions sont homogènes sur l'étendue d'un bâtiment, le géoréférencement d'une planche dépend donc fortement de la précision planimétrique de quelques bâtiments. Avec une transformation à effet local fort tel que la polynomiale d'ordre 3, les planches contenant un seul bâti ont des erreurs résiduelles très faibles, tandis que celles contenant quelques bâtiments dont l'erreur de position varie conservent une erreur plus élevée. On ne retrouve pas cet effet sur les points des rues, répartis de façon plus homogène.

Dans le cas du carroyage, la RMSE moyenne est particulièrement forte lorsque l'on utilise une transformation de Helmert, mais retombe dès l'utilisation d'une transformation affine avec plus de paramètres. Par rapport à Helmert, une transformation affine à 6 paramètres ajoute une rotation et un changement d'échelle (dans chaque direction). La différence de résidus entre ces deux transformations indique que les planches sont déformées localement. C'est d'ailleurs ce que l'on observe avec les transformations suivantes, à l'exception de la polynomiale d'ordre 3 : la précision et l'exactitude du géoréférencement augmentent. Le fait que la qualité du géoréférencement se dégrade pour des polynômes de degré supérieur à 2 relève également du sur-ajustement. En effet, il n'y a que 12 points d'amer par planche pour le cas du carroyage et l'estimation des paramètres d'un polynôme de degré 3 nécessite 10 points : il ne reste donc que deux degrés de liberté.

L'utilisation des points d'amer sur le carroyage et d'une transformation polynomiale d'ordre 2 semble un bon compromis entre qualité du géoréférencement et répartition des points. En effet, on a dans ce cas une RMSE moyenne sur l'ensemble de l'atlas de 1,12 mètre, tandis que la plus faible est de 0,47 mètres en utilisant les points des rues. Étant donné la qualité des tracés de l'atlas, une erreur de 65 centimètres est négligeable. De plus, cette erreur est particulièrement stable sur l'ensemble de l'atlas avec 30 centimètres d'écart type. Géoréférencé de cette façon, l'atlas présente donc des erreurs résiduelles réparties de manière plus homogène. Enfin, en utilisant un polynôme d'ordre 2, l'intégralité des planches de l'atlas peuvent être géoréférencées.

Répartition spatiale des erreurs résiduelles

Il est également intéressant d'observer la répartition spatiale des erreurs rési-

duelles pour plusieurs raisons. D'abord, on peut vérifier si cette répartition fait apparaître des structures dans les déformations du plan : décalage global, étirement, etc. D'autre part, elle permet d'observer si des phénomènes liés au levé topographique sont visibles. Par exemple, on peut s'attendre à ce que les erreurs soient plus faibles lorsque l'on est proche de la méridienne de France, de l'Observatoire Royal ou d'un point de triangulation majeur. La figure 3.28 montre la répartition spatiale de ces erreurs. Le tracé noir indique la limite des planches contenant des informations géographiques. La méridienne est figurée par une ligne rouge, et l'emplacement de l'Observatoire Royal par un point rouge. Il apparaît tout d'abord que

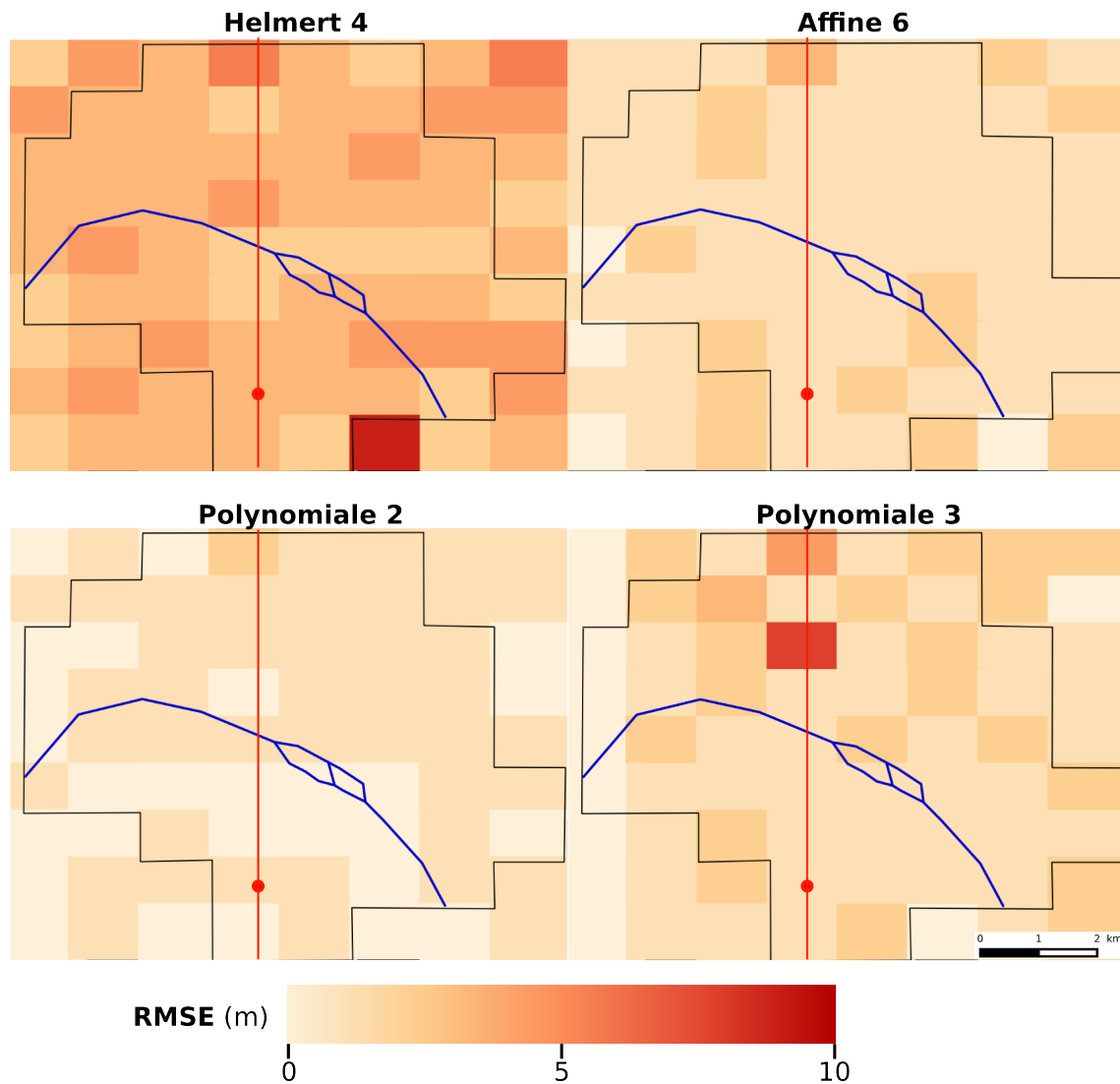


FIGURE 3.28 – Répartition spatiale des RMSE sur les planches de l'atlas de Verniquet dans le cas où les points d'amer sont placés sur le carroyage.

la répartition des erreurs quadratiques moyennes des planches ne semble décrire aucun motif particulier, ce qui est rassurant puisque tous les points du carroyage ont la même origine. Plus encore, aucune spécificité liée à la présence de la méridienne ou de l'observatoire n'est visible. Seule le cas de la transformation polynomiale d'ordre

3 fait montre une concentration des erreurs autour de la méridienne dans la partie nord de la ville. Cependant, on ne retrouve d'erreurs importantes que pour deux planches (4,12 et 7,14 mètres), ce qui ne permet pas de tirer de conclusions pour l'ensemble de l'atlas.

Choix d'un ensemble de points d'amer et d'une transformation

À la lecture des paragraphes précédents, la qualité du géoréférencement lorsque les points du carroyage sont utilisés avec une transformation polynomiale d'ordre 2 ressort clairement. La faible perte d'exactitude est compensée par une couverture totale de l'atlas et une forte homogénéité des erreurs résiduelles. On obtient alors un géoréférencement dont l'erreur quadratique moyenne sur l'ensemble de l'atlas est de 1,1 mètres, son maximum de 2,36 mètres et son minimum de 0,5 mètres. Étant donné l'âge de l'atlas, on peut considérer qu'il s'agit d'un géoréférencement de bonne qualité. Notons toutefois que ces mesures sont relatives au référentiel géographique utilisé, et non pas à la réalité. Les données du plan parcellaire de la ville de Paris étant précises à 20 centimètres, il en va de même pour nos mesures qui se trouvent donc entachées d'une imprécision du même ordre de grandeur.

Avantage du carroyage : continuité inter-planches

L'utilisation de points d'amer sur le carroyage du plan a l'avantage majeur de maximiser la correspondance des observations géohistoriques entre planches à condition que celles-ci subissent les mêmes déformations. Si cela permet d'obtenir un plan global en une seule partie à partir de l'atlas, il s'agit surtout de faire correspondre des observations géohistoriques réparties sur plusieurs planches (en particulier les rues) dans le but de faciliter la tâche de vectorisation. En effet, si les planches ne correspondent pas, la vectorisation du réseau routier nécessite la mise en place de spécifications de saisies permettant d'assurer la continuité des objets linéaires. La méthode la plus simple est de tracer des morceaux de rues fictives liant les parties d'une rue tracées sur deux planches. Une autre méthode consiste à "forcer" la correspondance entre les objets linéaires entre deux planches, par exemple en interpolant leur position. Ces méthodes présentent cependant l'inconvénient majeur de briser la correspondance entre les objets vectorisés et la carte elle-même.

3.5.4 Vérification du placement de la perpendiculaire et de l'approximation du système de projection de l'atlas par le Lambert I NTF

Nous disposons grâce au support géodésique qu'est le carroyage d'un géoréférencement particulièrement exact pour une carte ancienne. Non seulement l'erreur quadratique moyenne sur les planches de l'atlas est faible, mais elle est de plus particulièrement homogène avec une variation de 1.86 mètres. Toutefois, toute notre approche de géoréférencement sur carroyage n'est valide qu'à la seule condition que notre hypothèse sur l'approximation du système de projection par le Lambert I soit acceptable. Or, nous avons posé cette hypothèse sans la vérifier jusqu'ici. Dans cette section, nous proposons une méthode permettant de vérifier cette hypothèse en nous appuyant sur le géoréférencement. Nous présentons d'abord notre méthodologie de manière globale, puis dans le cas spécifique de vérification de notre hypothèse avant

de l'appliquer aux résultats du géoréférencement de l'atlas de Verniquet. Enfin, nous concluons sur la validité de notre hypothèse initiale.

Méthodologie globale : mesure des distorsions du plan

Le principe de notre méthodologie, déjà appliqué par (Balletti, 2006) repose sur l'utilisation de la transformation spatiale estimée à partir du carroyage pour transformer un second ensemble de points d'amer servant de points de mesure des distorsions du plan. Pour la suite, on parlera de points d'amer pour désigner les points servant à l'estimation de la transformation spatiale \hat{F} , et de **points de mesure** des points placés sur des objets géographiques dans le plan et le référentiel, transformés par \hat{F} après estimation (voir la figure 3.29). De la même façon que l'on note (p, p') un couple de points d'amer sur des éléments homologues dans la source et le référentiel, on écrit (m, m') l'équivalent pour les points de mesure. L'idée est

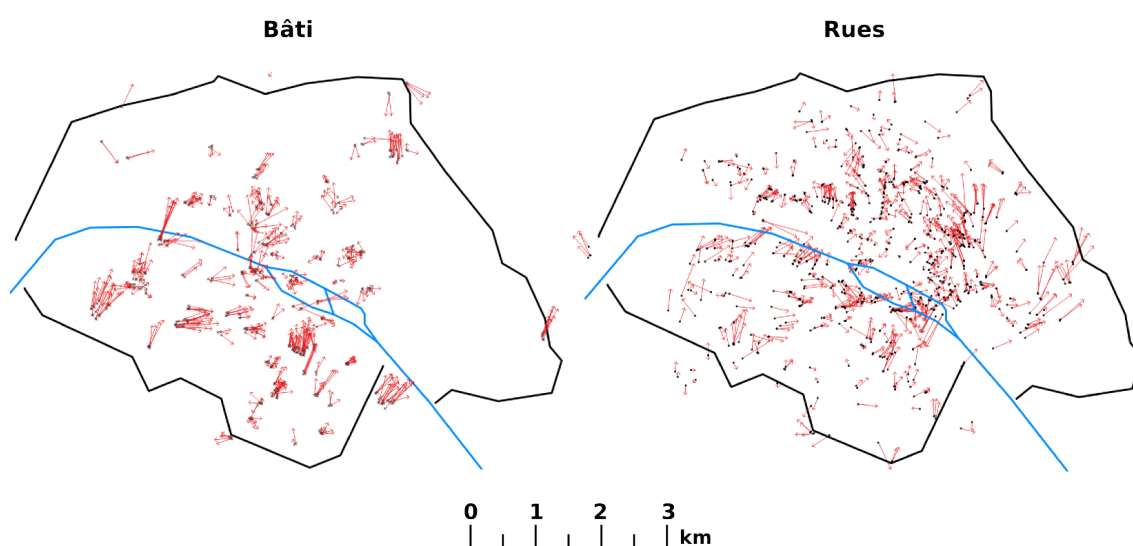


FIGURE 3.29 – Vecteurs de distorsion calculés pour les points de mesures *bâti* et *rues*. Vecteurs exagérés 100 fois.

la suivante : si l'on considère que la transformation estimée pour le carroyage est très proche de la véritable transformation F ($\hat{F}(p) \approx p'$, alors les éventuels écarts relevés entre un point de mesure m' dans le référentiel et la transformation de son correspondant dans la carte $\hat{F}(m) = n$ décrit la distorsion locale du plan. Le vecteur $\overrightarrow{m\hat{n}}$ est alors le **vecteur de distorsion** du plan au points m' . Ceci n'est cependant vrai qu'à condition que les points de mesures soient fiables.

Étant donné les résultats obtenus avec le carroyage sur un plan datant du XVIII^e siècle, nous considérons ici que $\hat{F} \approx F$. Puisque nous avons placé des points sur les rues et les bâtiments de l'atlas, nous pouvons les réutiliser comme points de mesure. La figure 3.30 résume cette méthodologie.

Application pour la vérification de l'hypothèse du système de projection

Supposons que notre hypothèse est fautive, i.e le système de projection de l'atlas de Verniquet ne peut être approximé par le Lambert I NTF. Cette hypothèse pourrait être fautive en particulier si la méridienne n'est pas la même (erreur dans

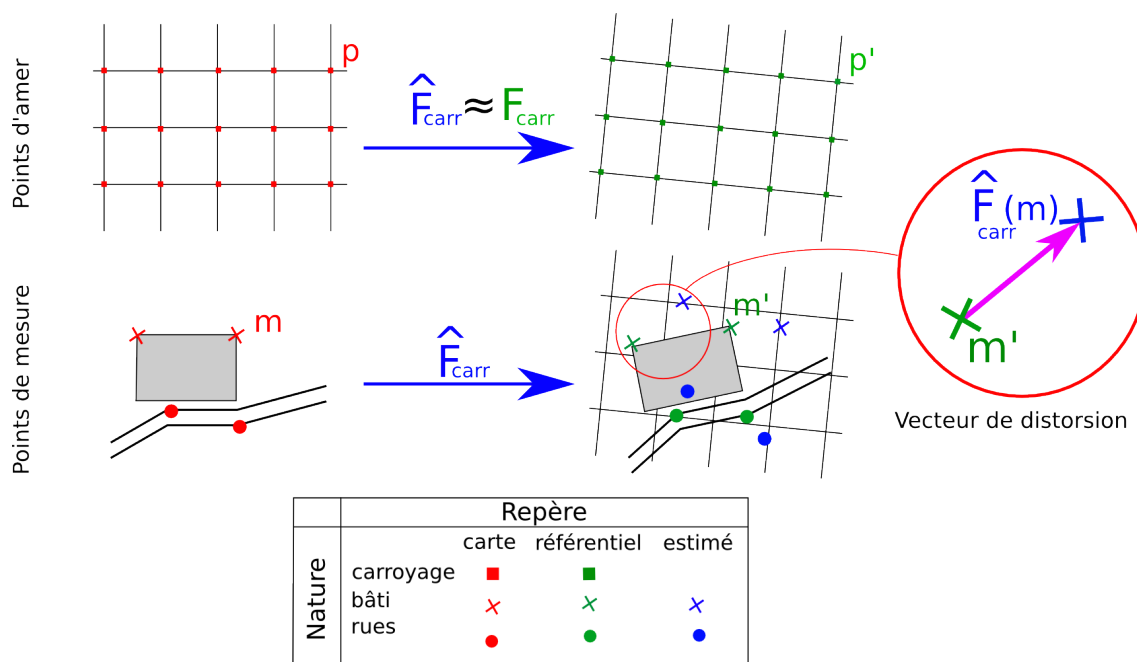


FIGURE 3.30 – Schéma illustratif de la méthodologie d'analyse du plan de Verniquet

les levés, re-calculs de sa position, etc.). Dans ce cas, le carroyage que nous avons créé dans le référentiel géographique, guidé par la méridienne actuelle, serait faux. Se pose également la question de la perpendiculaire. La façon dont nous l'avons estimée est empirique et nécessairement erronée. Tous ces éléments peuvent se cumuler, rendant faux le carroyage défini dans le référentiel. En conséquence, on peut avoir affaire à 3 types d'erreurs systématiques, résumées dans la figure 3.31 :

- Dans le plan, la méridienne et l'axe vertical du carroyage sont confondus (l'azimut du carroyage à l'Observatoire est nul). Sous notre hypothèse de base, le carroyage du plan dans le référentiel géographique et la méridienne sont également confondues. Or, s'il y a erreur, l'azimut à l'Observatoire peut ne plus être nul : le carroyage doit alors subir une rotation de centre l'Observatoire (cas 1 de la figure)
- Si la perpendiculaire a été mal placée, alors le plan subit une translation suivant l'axe nord-sud.
- Si l'axe de la méridienne est mal placé dans le plan (position fautive de l'Observatoire, décalage entre le contenu et le canevas géodésique, etc.), le plan subit cette fois une translation suivant l'axe est-ouest.

Si l'une de ces erreurs est présente, nous devrions la retrouver de manière systématique dans les distorsions mesurées sur l'ensemble de nos points de mesure. L'objectif est donc de vérifier s'il existe une erreur systématique horizontale, verticale ou d'orientation sur l'ensemble de l'atlas, mais également de mesurer ces erreurs. Ainsi, même si notre hypothèse d'approximation de la projection est fautive, nous connaissons alors les paramètres de la transformation permettant, pour Verniquet, de corriger cet écart, et nous pourrions déterminer son véritable système de projection.

Nous testons ces trois erreurs à partir des points de mesure définis sur les bâtiments

et les rues dans les sous-sections 3.5.4.1, 3.5.4.2 et 3.5.4.3. Dans chacun des trois cas, nous proposons une analyse visuelle à partir d'une cartographie des distorsions, accompagnée d'une seconde analyse s'appuyant sur un test d'autocorrélation spatiale. Ce type de test permet de mesurer la corrélation de mesures géographiques voisines. Dans notre cas, ce test d'autocorrélation nous permet d'explorer l'existence d'une erreur systématique. Nous nous appuyons plus particulièrement sur l'indice de corrélation de Moran (Moran, 1950), que nous détaillons brièvement dans le prochain paragraphe.

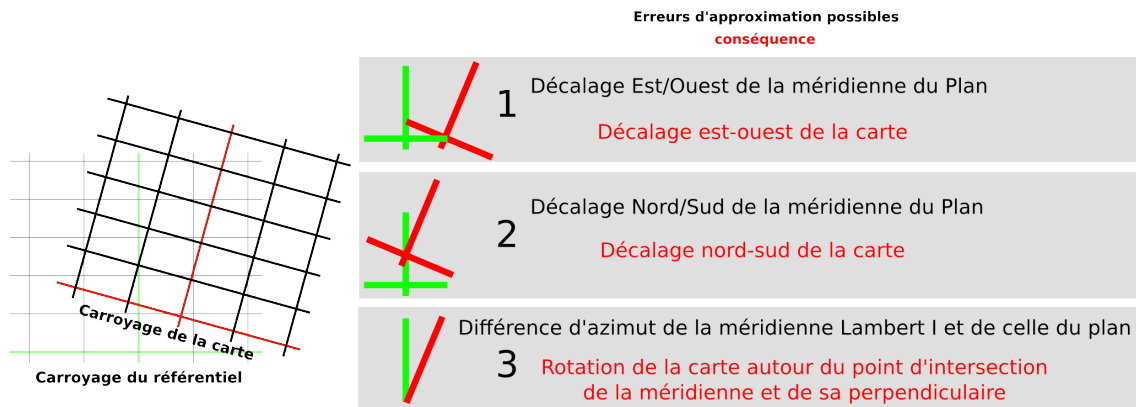


FIGURE 3.31 – Types d'erreurs systématiques provoquées par une erreur de définition du carroyage dans le référentiel géographique.

Note sur l'indice de Moran

L'indice de Moran (noté I) est une mesure d'autocorrélation spatiale décrivant la ressemblance entre des données voisines géographiquement. Son objectif est de donner une indication sur l'organisation spatiale des points dont la valeur se ressemble. L'indice est calculé pour un ensemble de classes de distances, et se lit au travers d'un corrélogramme (des exemples de corrélogrammes sont visibles dans les figures 3.33, 3.35 et 3.37). L'indice de Moran évolue sur l'intervalle $[-1; 1]$. Une autocorrélation positive pour une classe de distances C signifie que les distances situées à une distance C les unes des autres se ressemblent. À l'inverse, une autocorrélation négative indique que ces mesures sont différentes. Enfin, lorsque $I \approx 0$, la répartition des mesures est proche de l'aléatoire.

Dans notre cas, cet indice nous permet de vérifier s'il existe un décalage systématique de nos mesures verticalement, horizontalement ou s'il y a rotation. En effet, si le plan est entaché sur son ensemble d'une des trois erreurs décrites précédemment, nous devrions retrouver dans les distorsions mesurées une autocorrélation portant sur toute la surface de Paris.

Création des données nécessaires à l'analyse.

Les points de mesure sont trop peu nombreux (en particulier sur les bâtiments) pour qu'une carte interprétable soit produite, en particulier pour les bâtiments. Afin de faciliter l'exploration visuelle, nous créons une interpolation des vecteurs de distorsion calculés pour chaque point de mesure (bâtiments et rues) afin de créer un champ

de vecteurs dont nous pouvons cartographier les propriétés. En particulier, nous nous intéressons à la projection de ces vecteurs sur les axes verticaux et horizontaux du carroyage afin de déterminer les éventuelles erreurs systématiques correspondantes. Nous utilisons également son orientation par rapport à l'axe vertical pour explorer l'existence d'une rotation du carroyage. Nous avons utilisé une interpolation par distance inverse à une puissance ($p=2$) afin de produire un champ continu tout en conservant les distorsions locales. La norme des vecteurs de distorsion est ici exagérée 100 fois pour la visualisation). Ces champs de vecteurs interpolés sont utilisés uniquement pour les visualisations cartographiques. Pour les indices d'autocorrélation, seul les mesures brutes non interpolées sont utilisées pour ne pas introduire de corrélation artificielle dans les résultats. Les points de mesure sur les bâtiments et les rues étant spatialement organisés différemment, il faut prendre en compte les effets de cette organisation dans l'analyse. Les mesures sur les rues étant réparties de façon plus homogène dans le plan, le test d'autocorrélation sur les rues est considéré plus fiable.

3.5.4.1 Exploration de la présence d'une erreur systématique due à la perpendiculaire à la méridienne

La première erreur systématique à laquelle nous nous intéressons concerne le placement de la perpendiculaire à la méridienne de France dans le référentiel géographique. En cas d'erreur sur sa localisation, les bâtiments et les rues se retrouvent distordues suivant l'axe vertical : vers le nord si la perpendiculaire à été placée trop au nord de sa position réelle et vers le sud dans le cas contraire.

Pour mesurer cette erreur, nous étudions la projection des vecteurs de distorsion sur l'axe vertical de notre carroyage.

Cartographie des distorsions verticales

Les cartes de ces distorsions verticales, mesurées sur l'ensemble des points du bâti et des rues et interpolées, est présenté dans la figure 3.32. Les distorsions sont positives -en vert- lorsqu'elles sont dirigées vers le nord, et négatives -en rouge- vers le sud. La figure montre aussi les histogrammes des deux cartes. Précisons également que les deux cartes présentent uniquement la zone de l'histogramme surlignée en couleur correspondant afin d'éliminer les valeurs extrêmes rares gênant la visualisation. Chaque carte contient enfin les points de mesure correspondant au thème cartographique, et le tracé en noir des limites de Paris à la fin du XVIII^e siècle. On remarque tout d'abord que les distorsions sur majoritairement dirigées vers le nord, avec 97% de distorsions positives pour le bâti et 91% pour les rues. Sur les mesures non interpolées, on retrouve 74% de distorsions positives pour le bâti (avec une moyenne des distorsions positives de +1.9m) et 70% pour les rues (moyenne = +1.6m). La répartition spatiale de ces distorsions varie toutefois fortement pour un même thème cartographique et entre eux. Sur le bâti, les distorsions sont constantes pour un même bâtiment mais peuvent varier largement d'un bâti à l'autre. Par exemple, le palais de justice sur l'île de la Cité subit une distorsion vers le sud de 1 mètre, tandis que le Louvre est décalé de 4 mètres au nord. Des distorsions positives côtoient donc des distorsions négatives : il ne semble pas y avoir de schéma général qui se dégage pour le thème du bâti, mais plutôt des distorsions propres à

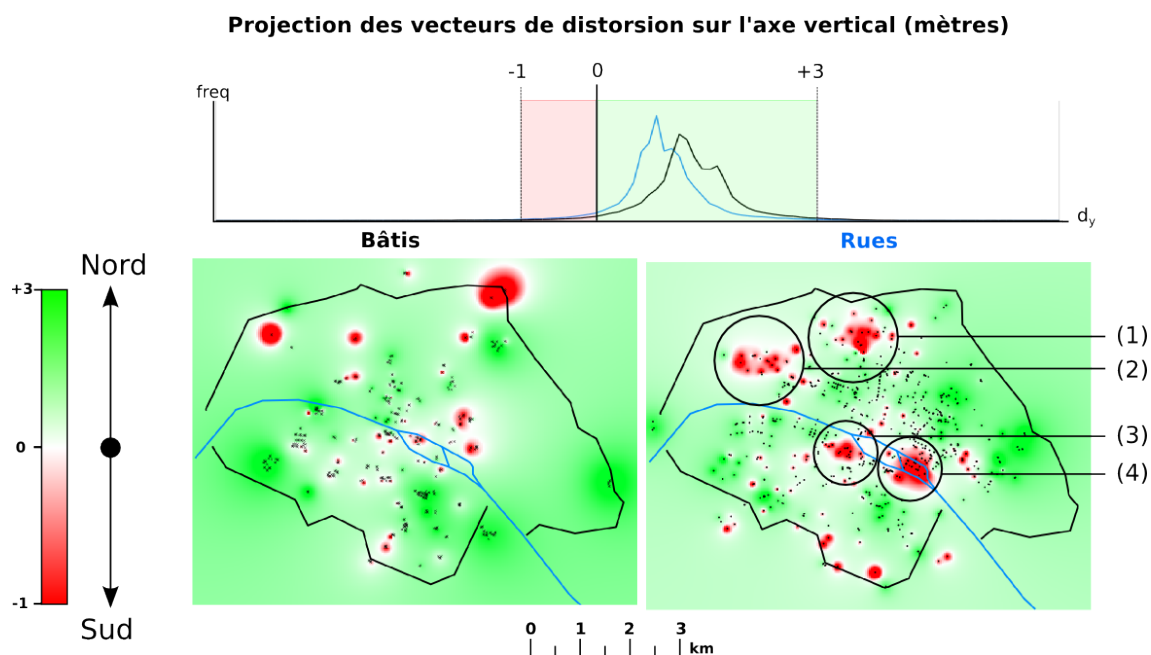


FIGURE 3.32 – Carte interpolée des distorsions verticales mesurées sur le bâti et les rues de l'atlas de Verniquet.

chaque bâtiment. Nous explorerons cette hypothèse plus tard. Pour les rues cette fois, quatre zones se différencient du reste du plan, plutôt distordu vers le nord. La première (indiquée (1)) correspond aux alentours de Notre-Dame-de-Lorette, dans l'actuel quartier du Faubourg-Montmartre, qui subit une distorsion verticale allant jusqu'à -4 mètres. La seconde est ciblée autour de l'actuel palais de l'Élysée (2), avec -0.9 à -3m de distorsion. Enfin les deux dernières zones sont centrées respectivement sur le Collège de Mazarin (actuel Institut de France, (3)) et l'île Saint-Louis (4). Étant donné la répartition des distorsions sur les rues, nous pouvons envisager l'existence d'une distorsion systématique vers le nord, avec cependant certaines zones délimitées distordues dans le sens opposé en raison d'erreurs de levé topographiques ciblées. Pour creuser cette hypothèse et renforcer cette première analyse visuelle, nous appliquons un test d'autocorrélation spatiale.

Autocorrélation des distorsions verticales

Le corrélogramme de l'indice de Moran sur les distorsions verticales des données brutes est présenté dans la figure 3.33. L'indice est calculé pour chaque classe de distance couvrant jusqu'à toute la largeur de Paris. Afin de vérifier la significativité statistique de ces indices, nous avons effectué pour chacun un test de Khi-deux (Greenwood et Nikulin, 1996). Les indices statistiquement significatifs sont indiqués par des points rouges³². La courbe d'évolution de l'indice de Moran pour le bâti est tracé en vert, et celui des rues en or. Il existe une corrélation des distorsions en dessous de 300 mètres ($I=0.6$), qu'il s'agisse du bâti ou des rues. Pour les bâtiments, cette corrélation s'explique par le fait que les points de mesures placés

32. Il s'agit de ceux pour lesquels l'hypothèse selon laquelle les valeurs sont le fruit du hasard peut être rejetée avec un degré de confiance supérieur à 95% .

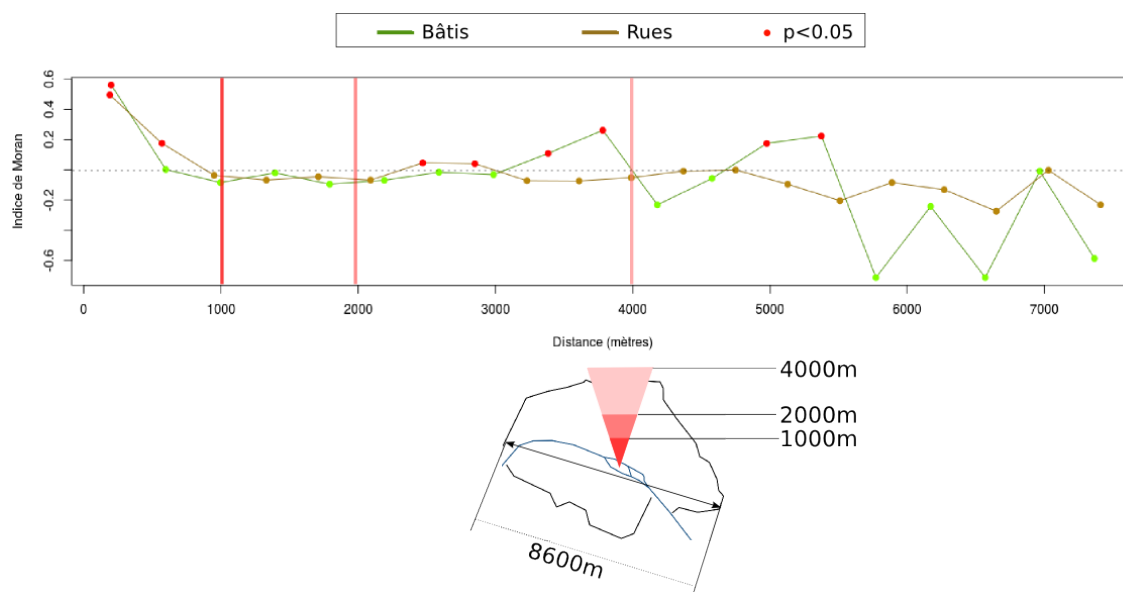


FIGURE 3.33 – Corrélogramme traçant l'indice de Moran pour 19 classes de distances sur l'atlas de Verniquet, pour les distorsions verticales sur le bâti (en vert) et sur les rues (en or). Les points significatifs ($p < 0.05$) sont en rouge.

sur ceux-ci sont regroupés et leurs distorsions cohérentes pour un même bâtiment (autrement dit, le bâtiment subit une transformation rigide : il est mal placé mais géométriquement juste). Pour les rues, elle peut s'expliquer de façon similaire : les levés topographiques sont cohérents pour une même rue ou des rues proches³³.

Toutefois, le corrélogramme ne montre pas de décalage systématique nord ou sud. En effet, si tel était le cas, l'autocorrélation des distorsions pour les grandes distances devrait être positive³⁴. Or, l'indice de Moran pour les rues est proche de 0 jusqu'à 5000 mètres, puis une faible autocorrélation négative apparaît, non significative. Dans le cas des rues, la disposition des distorsions est donc quasi-aléatoire jusqu'à des distances proches de 5000 mètres. Même au delà, la non-significativité des mesures ne permet pas d'exclure une disposition aléatoire.

Pour le bâti, l'indice augmente ($I \approx 0.2$) pour les distances de 3800 mètres et 5000-5400 mètres, qui correspond à l'éloignement des bâtiments situés rive gauche et rive droite. Ces points étant significatifs, il existe une faible corrélation des distorsions des bâtiments. Si l'on mesure les distances entre mesures, on peut identifier la direction des distorsions pour ces classes de distances. Ainsi, il s'agit à 3800 mètres d'une corrélation entre distorsions négatives, et à 5000-5400 mètres des distorsions positives.

Conclusion sur l'erreur de placement de la perpendiculaire à la méridienne

Les cartographiques et le test d'autocorrélation ne permet donc pas d'identifier une distorsion systématique sur l'axe vertical. Cependant, l'existence de lieux distordus dans un sens opposé de celui, majoritairement visible sur les cartographies, du

33. Les levés des rues étaient en effet faits rues par rue, de proche en proche.

34. Au delà de quelques centaines de mètres, les effets locaux des levés n'existent plus.

nord, dégrade les résultats d'un tel test d'autocorrélation. En effet, il est envisageable que cette distorsion systématique existe, mais qu'elle soit localement compensée par des erreurs dans le plan lui-même. Nous avons en outre vu que l'erreur moyenne vers le nord est d'environ 1.8 mètres sur les deux thèmes. Il est cependant nécessaire de vérifier l'existence des deux autres erreurs systématiques avant de conclure.

3.5.4.2 Exploration de la présence d'une erreur de positionnement de l'axe de la méridienne dans l'atlas

Nous avons ensuite effectué la même étude sur la projection des vecteurs de distorsion sur l'axe horizontal. L'objectif est cette fois de vérifier s'il existe un décalage est-ouest du carroyage de Verniquet dans le référentiel géographique qui indiquerait une erreur de positionnement du méridien de Paris (Lambert I) vers l'ouest ou l'est de la méridienne.

Cartographie des distorsions horizontales

Les distorsions mesurées sont illustrées dans la figure 3.34. Le profil de ces dis-

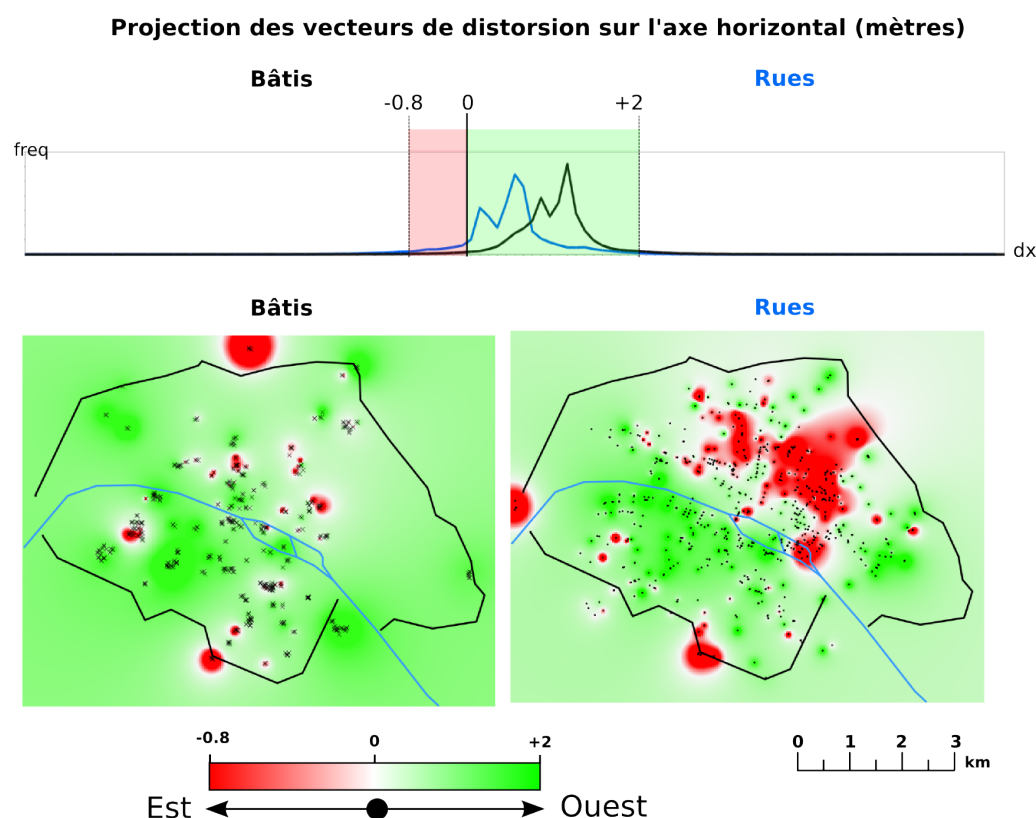


FIGURE 3.34 – Carte interpolée des distorsions horizontales mesurées sur le bâti et les rues de l'atlas de Verniquet.

torsions est relativement similaire à la précédente carte : la majorité des distorsions sont dirigées vers l'ouest, bien qu'elles évoluent dans un intervalle plus restreint. En effet, 98% des mesures se situent dans l'intervalle $[-0.8; +2]$ mètres. L'effet est particulièrement marqué pour les points de mesure placés sur le bâti avec 76% de

distorsions positives sur les mesures non interpolées et une moyenne de +1.18 mètres. Sur les rues, l'effet est moins marqué : seules 59% des distorsions sont dirigées vers l'ouest, avec une moyenne de seulement 80 centimètres. On constate toutefois, pour les rues toujours, une organisation spatiale des distorsions tout à fait singulière avec un décalage vers l'est dans toute le centre ancien de la rive droite, et à l'ouest pour le reste à l'exception de quelques points isolés et des barrières de Sainte-Marie à l'ouest et d'Enfer au sud de la ville.

À l'inverse des distorsions verticales pour lesquelles les cartographies tendaient à montrer un décalage systématique vers le nord, il est plus difficile de l'affirmer ici pour l'ouest en raison de la différence marquée entre rives gauche et droite de la Seine.

Corrélogramme des distorsions horizontales

Le corrélogramme des distorsions horizontales, visible dans la figure 3.35 ressemble également au précédent. On retrouve une corrélation significative aux faibles

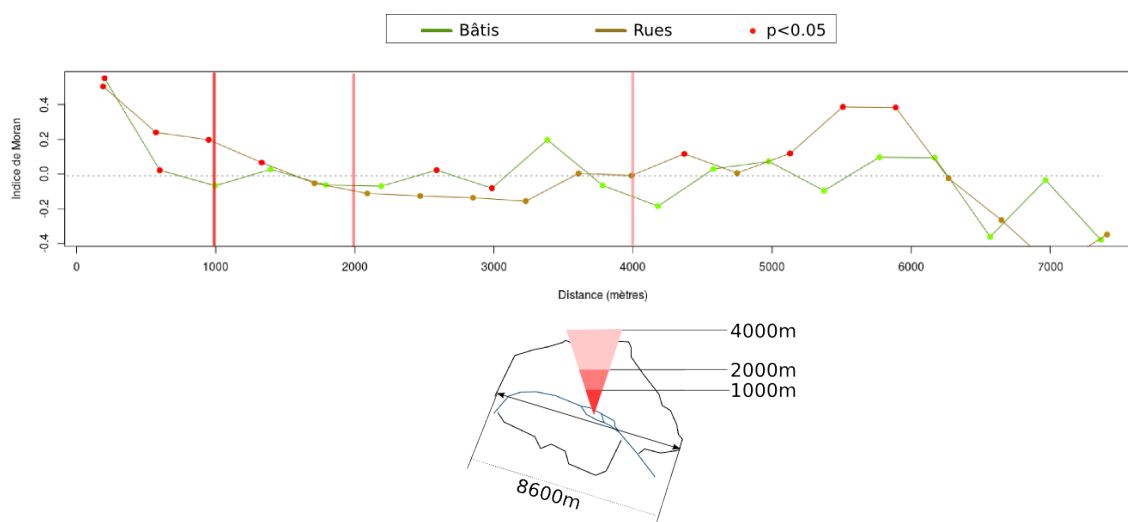


FIGURE 3.35 – Corrélogramme traçant l'indice de Moran pour 19 classes de distances sur l'atlas de Verniquet, pour les distorsions horizontales sur le bâti (en vert) et sur les rues (en or). Les points significatifs ($p < 0.05$) sont en rouge.

distances qui indique une cohérence dans les distorsions subies par des groupes de bâtis ou de rues voisins. Les variations sur les bâtiments sont toutefois plus importantes, ce qui provoque une chute plus rapide de l'indice : dès 500 mètres, il n'y a plus de corrélation entre mesures sur le bâti. Aucune corrélation significative n'est ensuite visible avant les grandes distances, à partir de 5000 mètres. On retrouve de nouveau le même effet que précédemment, mais cette fois sur les rues : des groupes de distorsions positives ou négatives se trouvent éloignées de plus de 5 kilomètres, avec entre elles des zones de distorsion variées. Au delà des 6000 mètres, il n'y a plus que quelques points de mesures situées aux extrêmes est et ouest, dont les distorsions sont opposées. Étant donné la surface de Paris, une distorsion horizontale systématique serait visible principalement pour des distance allant de 0 à environ 5000 mètres. Or, le corrélogramme ne montre de nouveau aucune tendance nette à corrélation. Au contraire même, il fait ressortir une faible autocorrélation négative

autour de 3000 mètres sur les rues : il s'agit de l'opposition rive gauche-rive droite visible sur la carte précédente.

Conclusion sur l'erreur de placement de la méridienne du plan dans le référentiel géographique

Nous avons montré qu'il n'existe pas de décalage horizontal systématique significatif sur les mesures effectuées sur les bâtiments comme sur les rues. Nous pouvons donc considérer que le point d'origine à l'Observatoire est le même pour le méridien de Paris (Lambert I) que pour celui du plan placé dans le référentiel géographique. Il faut toutefois rester prudent. En effet, nous considérons ici que les erreurs résiduelles issues du géoréférencement par carroyage sont négligeables. Or, nous traitons ici de distorsions du même ordre de grandeur que les erreurs résiduelles du carroyage. En effet, erreurs et distorsions sont particulièrement faibles, de l'ordre du mètre. Nos mesures peuvent donc être biaisées par ces erreurs résiduelles.

3.5.4.3 Exploration de la présence d'une erreur d'orientation du carroyage

Nous testons enfin l'hypothèse d'une rotation du carroyage du plan. Une telle rotation peut apparaître si l'azimut du carroyage du plan dans le référentiel géographique n'est pas nul. Autrement dit, il y a rotation si la méridienne du plan n'est pas réellement orientée vers le nord géographique.

Pour tester cette hypothèse, nous mesurons l'angle entre les vecteurs de distorsion issus des points de mesure et l'axe horizontal dans le référentiel géographique. En effet, si le carroyage du plan dans le référentiel doit subir une rotation vers la gauche, centrée sur le point d'origine de la méridienne à l'Observatoire, nous devrions retrouver une répartition des orientations en deux groupes opposés au nord et au sud de Paris.

Cartographie des orientations des vecteurs de distorsion

Les résultats des mesures d'orientation apparaissent dans la figure 3.36. La variable visualisée correspond au cosinus de l'angle formé par le vecteur d'orientation avec l'axe horizontal. Une valeur de 1 indique donc une orientation vers l'ouest et -1 vers l'est. Puisqu'il s'agit d'un cosinus, les orientations nord et sud ne sont pas distinguées. La méridienne et sa perpendiculaire sont figurées en vert.

On retrouve dans ces figures les types de distorsion que nous avons déjà montrés dans les cartes de distorsion verticales et horizontales. Pour les bâtiments, nous retrouvons les variations de distorsions entre eux. Par exemples, l'école militaire (en (1)) est distordue dans une direction pratiquement opposée à l'hôpital des incurables (en (2)), situé 900 mètres au sud-est. On retrouve le même phénomène aux environs de Sainte-Geneviève (l'actuel Panthéon, (3)) ou du palais cardinal (4).

La carte des rues présente exactement le même profil de distorsion que ce que nous avons noté pour les distorsions horizontales. Tout le centre de la rive droite au delà de la rue de Rivoli et jusqu'au delà des boulevards Saint-Martin et Saint-Denis est à l'opposé du reste de Paris. Ici encore, les deux barrières Sainte-Marie (5) et d'Enfer (6) et la zone couvrant l'île Saint-Louis et Saint-Paul (7) se détachent de leur voisinage.

Orientation des vecteurs de distorsion par rapport à l'axe horizontal

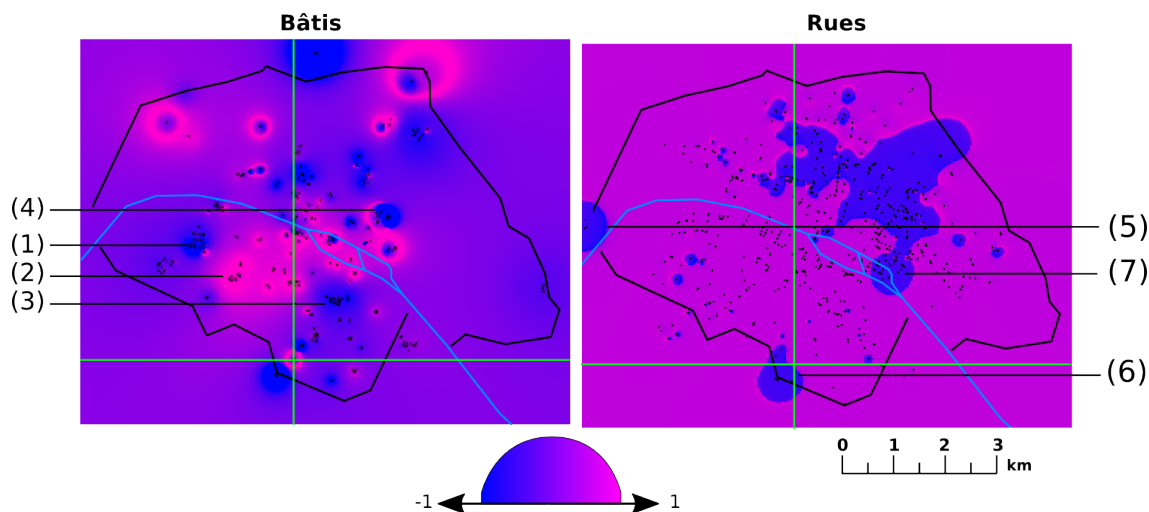


FIGURE 3.36 – Orientation des vecteurs de distorsions sur les bâtis et les rues par rapport à l'axe horizontal.

Visuellement, les deux cartes ne montrent pas clairement le profil que l'on recherche : une distorsion dans un sens au nord, et dans l'autre au sud. En effet, bien que deux zones opposées se détachent sur le plan des rues, la rotation du carroyage devrait induire des orientations similaires dans toute la partie nord et ouest de la rive droite (et à l'opposé dans la partie sud-est de la ville). De plus, les distorsions des bâtiments suivent une répartition qui ne présente pas d'organisation nettement visible.

Corrélogramme des orientations des vecteurs de distorsions

Le corrélogramme des orientations est visible dans la figure 3.37. Cette fois, à l'inverse des erreurs de translation, nous recherchons une division de l'espace en deux blocs et donc une autocorrélation positive aux faibles distances et négative aux distances de l'ordre du rayon de la ville (~4000m). On retrouve une autocorrélation positive de 0.2 pour les rues à 0.4 pour le bâti aux faibles distances (quelques centaines de mètres), ce qui correspond aux cartes précédentes. On ne retrouve cependant aucune autocorrélation aux alentours de 4000 mètres.

Conclusion sur l'erreur systématique liée à la rotation du carroyage

Étant donné les mesures visibles sur les cartes d'orientation et l'absence de corrélation négative aux environs du rayon de Paris, nous rejetons l'hypothèse d'une rotation du carroyage.

3.5.4.4 Conclusion sur l'exactitude du carroyage de l'atlas de Verniquet

L'absence d'erreur systématique sur la position horizontale et la rotation du carroyage indique que les systèmes de projection de l'atlas et le système Lambert I NTF sont suffisamment proches pour que les éventuelles erreurs induites par leurs différences soient inférieures aux distorsions propres aux feuilles de l'atlas.

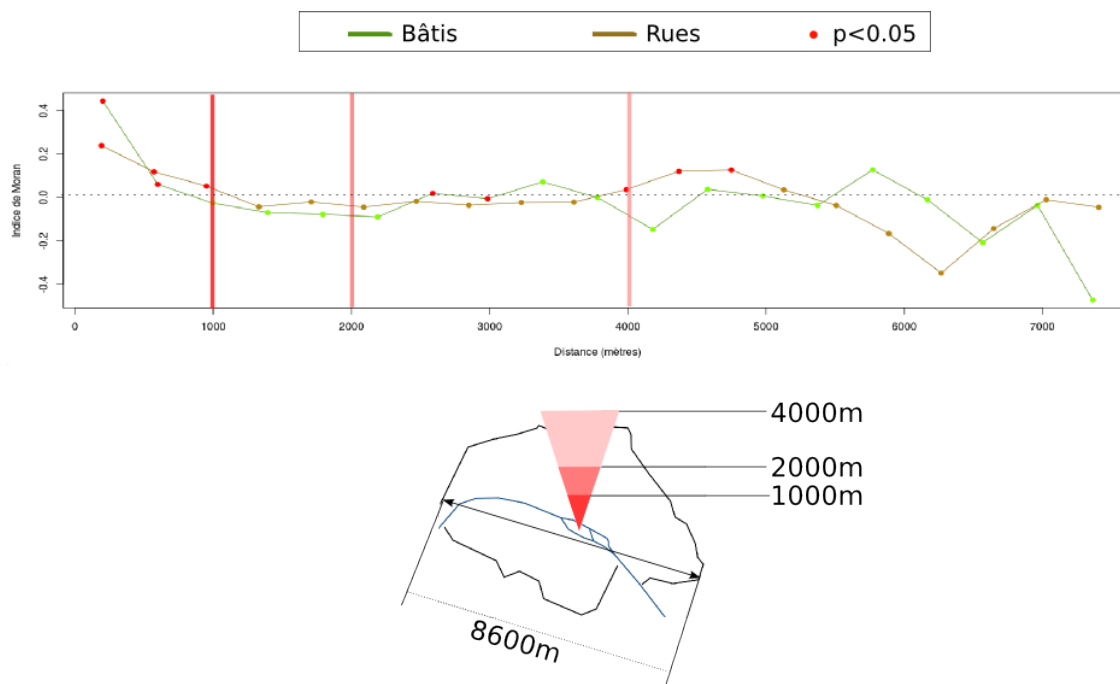


FIGURE 3.37 – Corrélogramme traçant l'indice de Moran pour 19 classes de distances sur l'atlas de Verniquet, pour les orientations des vecteurs de distorsion sur le bâti (en vert) et sur les rues (en or). Les points significatifs ($p < 0.05$) sont en rouge.

L'étude du décalage vertical nous a permis quant à lui de faire ressortir une distorsion vers 1.8 mètres vers le nord sur une majorité des points de mesure, ce qui indique une erreur de positionnement de la perpendiculaire à la méridienne dans la même direction. Nous n'avons toutefois pas effectué de nouveau géoréférencement afin de corriger cette erreur. Cette opération demanderait également une nouvelle validation, en particulier pour les zones dont les distorsions sont dirigées vers le sud, qui se trouveraient alors encore augmentées.

Les différentes cartographies proposées permettent d'aller au delà de l'objectif fixé pour cette section et d'étudier la qualité du contenu du plan pour les deux thèmes cartographiques des rues et des bâtiments. En effet, au travers des mesures de distorsion des bâtis et des rues, on voit apparaître les distorsions dépendantes des opérations de levé topographiques et de dessin.

Nous proposons, dans la section suivante, d'étudier plus en profondeur ces distorsions. D'abord, la connaissance des distorsions de chaque thème cartographique offre la possibilité de dresser un premier modèle des distorsions de ce thème. Ce modèle peut ensuite servir à corriger localement les objets vectorisés extraits des plans. Il fournit également des métadonnées de précision planimétrique locale propre à chaque thème cartographique.

3.5.5 Analyse des distorsions des thèmes "bâti" et "rues" de l'atlas

Les points de mesures placés sur deux thèmes cartographiques distincts donnent l'occasion de les confronter afin de critiquer la qualité géométrique du contenu de l'atlas. Cette critique peut s'effectuer sur deux points :

- l'étude des distorsions de chaque thème par rapport au carroyage géoréférencé (dont l'erreur d'estimation est jugée négligeable) permet d'explorer la qualité du levé topographique et du report sur le plan des rues et du bâti de façon indépendante. Ceci permet également de tracer de façon qualitative une première approximation du modèle de distorsion de chaque thème. Un tel modèle peut alors être utilisé pour corriger les distorsions des éléments vectorisés de ce thème.
- l'étude comparée des distorsions entre les deux thèmes offre quant à elle la possibilité de vérifier si les différents éléments composant la carte sont cohérents entre eux. En effet, le géoréférencement considère habituellement que tous les objets géographiques représentés dans l'image source suivent le même modèle de distorsion. Or, cette hypothèse ne peut pas être posée avec certitude pour des plans anciens dont le levé est souvent long et chaotique. Dans certains cas, les différents thèmes cartographiques sont issus de levés différents. C'est notamment le cas de l'atlas de Verniquet pour lequel les bâtiments et les rues ont été levés séparément (voir le chapitre 2).

Nous traitons ces deux points ainsi : tout d'abord, nous étudions les vecteurs de distorsion de chacun des deux thèmes cartographiques "bâti" et "rues", puis nous regardons plus en détail les différences entre ces deux thèmes.

Hypothèse sur la répartition des distorsions et discussion sur sa validité

Les points de mesure sur les bâtiments et les rues sont spatialement répartis de manière différente et ne sont pas situés au même endroit. De plus, il y a nettement plus de mesures pour les rues que pour le bâti. Pour être capable de confronter les deux thèmes, nous nous appuyons sur l'interpolation définie précédemment. Cela revient à considérer que les points de mesure que nous avons placés sont suffisants pour fournir une bonne estimation du modèle de distorsion du thème correspondant. Si, pour les rues, nous pouvons considérer cette hypothèse valable (les points sont nombreux, répartis sur tout Paris), le cas des bâtiments est nettement plus problématique. En effet, nous avons vu que les distorsions de ce thème étaient liées à chaque bâtiment. Les résultats de cette section seront donc valides essentiellement dans les zones où les deux thèmes sont représentés par une forte densité de points de mesure, c'est à dire le centre ancien de Paris (à l'intérieur des boulevards).

3.5.5.1 Distorsions par thèmes

Lors de l'exploration des erreurs systématiques, nous avons pointé des distorsions spécifiques à chacun des thèmes cartographiques "bâti" et "rues". Nous explorons plus précisément ces spécificités ici.

Quantité de distorsion

Nous nous intéressons d'abord à la répartition spatiale de la quantité de distorsion pour chacun des thèmes. Celle-ci est illustrée par la figure 3.38. 98% des distorsions du bâti sont comprises entre 0.5 et 3.2 mètres, avec un maximum de 6.3 mètres pour la place du Trône (4). L'amplitude des distorsions est plus faible sur les rues, atteignant 4 mètres aux alentours de Sainte-Marguerite (6), et dont 98% des

valeurs sont inférieures à 2.5 mètres.

On voit de nouveau apparaître la forte variation des distorsions sur les bâtiments, en particulier pour l'école militaire, l'hôpital des Invalides (1) ou encore la Pitié Salpêtrière (2). Ainsi, l'école militaire présente une distorsion moyenne de 3 mètres, pour 0.5 mètres en moyenne aux Invalides, deux bâtiments éloignés de quelques centaines de mètres. On peut penser que l'on retrouve ici un effet de la construction même de l'atlas. En effet, les plans des bâtiments étaient réalisés de façon indépendante les uns des autres, et donc le fruit de levés topographiques différents. Toutefois, la partie du centre ancien aux alentours de l'Hôtel de Ville (3) reste homogène avec une distorsion faible.

Les distorsions des rues, quant à elles, se répartissent sur trois grandes zones. D'une part, la majorité de la rive gauche (7) et toute l'île de la Cité sont distordues de 1.5 mètres à 4 mètres. La rive droite est à l'inverse : moins distordue et moins homogène. L'axe correspondant aujourd'hui à la rue de Rivoli (5) est particulièrement visible, avec une distorsion presque nulle sur toute sa longueur. Enfin, une nouvelle fois, la rue de Charonne et ses environs (6) se détache de son voisinage, mais les distorsions restent plus faibles que dans la rive gauche (1 mètre à 2.5 mètres). Il est de nouveau possible que l'on retrouve un effet des levés topographiques du plan, ceux de la rive droite et de la rive gauche n'ayant pas été réalisés au même moment. Il serait intéressant de confronter ces mesures à reconstitution des opérations de levé à partir des minutes de l'atlas.

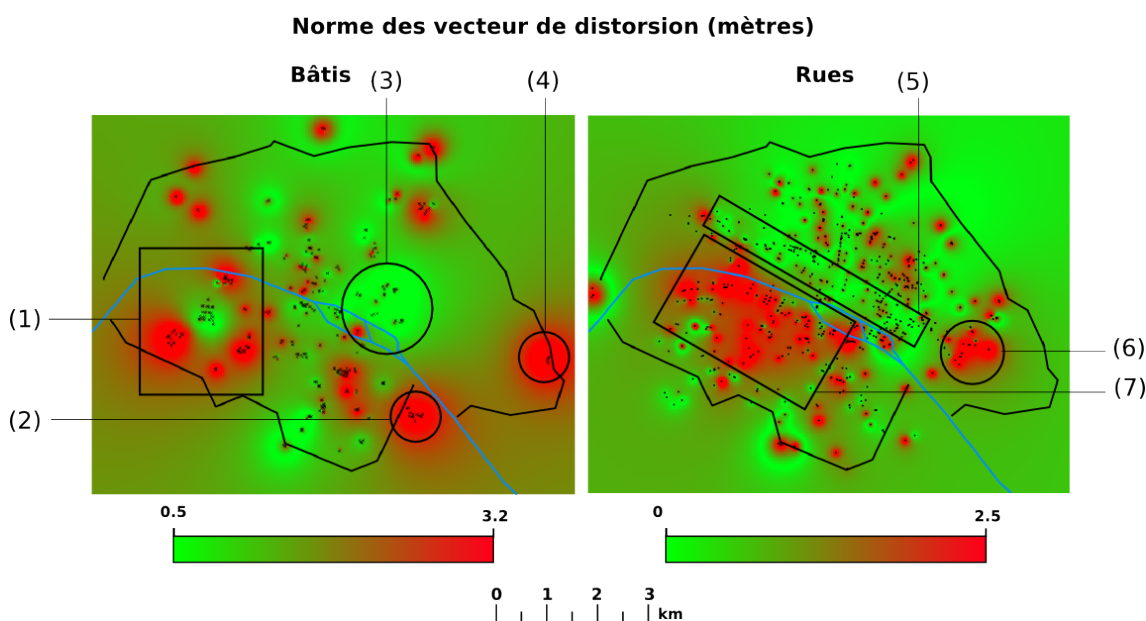


FIGURE 3.38 – Norme des vecteurs de distorsion pour les thèmes cartographiques "bâti" et "rues"

Orientation moyenne des distorsions

Afin de compléter ces premières remarques, nous avons également cartographié les orientations des vecteurs de distorsion. Ceux-ci nous renseignent en effet sur la direction dans laquelle chaque thème se trouve distordu. Couplés à l'analyse de la

quantité de distorsion des thèmes, ils permettent de tracer, de façon qualitative, un premier modèle des distorsions des différents thèmes cartographiques de l'atlas de Verniquet.

Cette cartographie, présentée dans la figure 3.39 est un lissage par moyenne glissante des orientations des vecteurs de distorsion. Ceux-ci sont ensuite classés selon 8 orientations. La carte des orientations sur le bâti présente toujours le même type de répartition, centrée sur chaque bâtiment. Deux zones d'orientations similaires se détachent. L'une (1), regroupe un ensemble de bâtiments distordus vers le nord-ouest : il s'agit du palais royal et de l'actuel Conservatoire National des Arts et Métiers. La seconde (2), distordue vers l'est, couvre un espace sur la rive gauche allant de Saint-Placide à l'île de la Cité, et poursuit jusqu'à la rive droite en englobant les abords est du Louvre et l'Hôtel de Ville. L'orientation des vecteurs de distorsion n'est cependant significative que lorsque la distorsion est importante. Or, la norme des vecteurs de distorsion varie largement au sein de ces deux zones : il n'est donc pas possible de se prononcer sur une structure globale des distorsions du bâti. Nous pouvons donc considérer qu'il n'existe pas de modèle général des distorsions du bâti, mais que celles-ci sont propres à chaque bâtiment composant la carte.

Sur les rues, nous retrouvons de nouveau la séparation en deux grandes zones que l'on pouvait déjà noter précédemment. Une première zone, composée du centre ancien au delà de la rue du faubourg Saint-Honoré (3), de l'île Saint-Louis et des environs de l'église Saint-Paul (4) est distordue vers le sud-ouest. Tout le reste de la ville est exactement à l'opposé, avec une orientation nord-est généralisée, à l'exception de quelques points et des barrières Sainte-Maire (5) et d'Enfer (6). L'axe de la rue du faubourg saint-Honoré lui-même suit la direction majoritaire vers le nord-est. Contrairement à celui du bâti, les distorsions du thème des rues semble donc s'organiser autour de deux grands ensembles : le centre ancien de la rive droite, assez peu distordu vers le sud-ouest s'oppose au reste du plan, composé de zones plus fortement distordues vers le nord-est. Les barrières font figure d'exception, mais elles sont situées dans des lieux très peu denses en points de mesure. Leur saillance est donc surtout le résultat de l'interpolation.

Vers un modèle des distorsions de l'atlas de Verniquet ?

Les analyses du plan que nous venons de présenter nous permettent de tirer deux conclusions sur la qualité du contenu de l'atlas de Verniquet. D'abord, avec 98% des mesures de distorsion inférieures à 3.2 mètres (pour 6.3 mètres maximum) pour le bâti et 2.5 mètres (5.4 mètres maximum) pour les rues, nous pouvons raisonnablement dire que les distorsions de ces thèmes sont faibles. Avec une erreur d'estimation du géoréférencement de 1.1 mètres sur le carroyage, nous obtenons donc une distorsion maximale pour 98% des points de mesure de 4.2 et 3.2 mètres. Une telle distorsion reste particulièrement faible pour un plan ancien : la promesse d'exactitude géométrique d'Edmé Verniquet semble belle et bien tenue.

De plus, les différentes cartographies que nous avons proposées nous ont permis de montrer les différences existant entre deux thèmes de l'atlas. En particulier, nous avons fait ressortir l'organisation des distorsions de chacun de ces thèmes : ciblées sur chaque bâti pour l'un, réparties en quelques grandes zones pour l'autre. Chaque thème semble donc obéir à sa logique propre et aucune véritable cohérence n'appar-

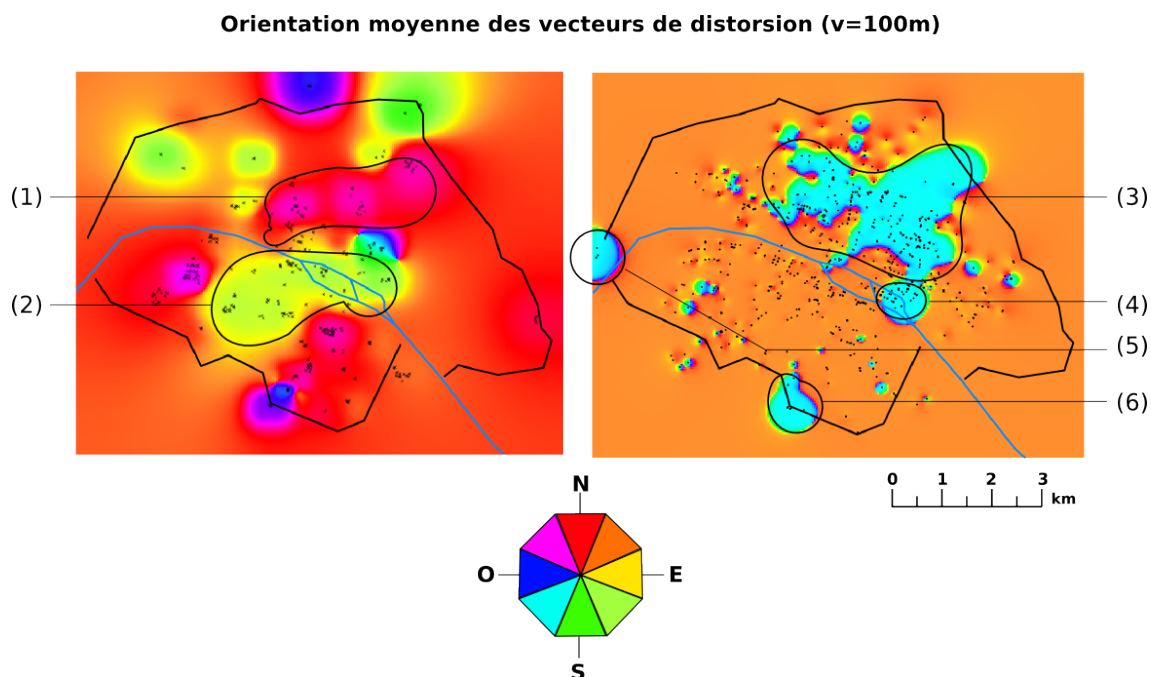


FIGURE 3.39 – Orientation lissée des vecteurs de distorsion par moyenne glissante dans une fenêtre de taille v .

rait entre eux. Ceci peut s'expliquer par le fait que bâti et rues ne sont pas issus des mêmes opérations de levé topographique et n'ont pas été reportés sur le plan au même moment (voir (Pronteau, 1986)). Nous retrouvons donc, grâce à ces points de mesure, certaines pratiques de construction d'une carte qui ont abouti à la création de l'atlas de Verniquet.

Finalement, ces différentes mesures constituent un premier pas vers la création d'un modèle des distorsions de l'atlas de Verniquet, thème par thème. Si l'on connaît celui-ci, il est alors possible de corriger le plan. De plus, si le contenu de l'atlas est vectorisé, il est alors possible de corriger les positions des objets extraits grâce à un tel modèle.

L'atlas de Verniquet, un référentiel géographique pour l'espace parisien ancien ?

Lorsque l'on géoréférence des plans de Paris antérieurs au XIX^e siècle ou que l'on souhaite géolocaliser des données dans l'espace ancien de la ville, il faut disposer d'un référentiel géographique sur lequel vont être projeté le plan ou les données. Or, ceci nécessite de pouvoir situer des éléments d'un espace ancien dans un référentiel récent. Avec les transformations importantes qu'a connu Paris après la Révolution, de nombreuses portions de l'espace parisien ont disparu, empêchant ainsi géoréférencement et géolocalisation. Il est donc intéressant de disposer d'un référentiel géographique sur cet espace pré-révolutionnaire, qui peut agir comme un pivot. Nous venons de montrer que l'atlas de Verniquet est une source cartographique sur l'espace de l'Ancien Régime dont la précision géométrique est remarquable. De plus, nous avons connaissance des distorsions que subissent chacun des thèmes principaux

du plan. Nous pensons donc que cet atlas géoréférencé peut être utilisé comme un référentiel géographique fiable pour la localisation de données anciennes ainsi que pour le géoréférencement de plans antérieurs à la Révolution. Puisque les distorsions de l'atlas et l'exactitude de son géoréférencement sont connues, il serait alors possible de connaître, par propagation d'erreurs, la précision planimétrique d'une donnée géolocalisée sur cet atlas par rapport au plan parcellaire de la ville actuelle.

3.5.6 Conclusion et géoréférencement des autres plans de Paris

Au travers de cette section, nous avons donc montré la pertinence de l'approche de géoréférencement proposée pour les plans topographiques anciens de Paris. En nous appuyant sur le carroyage, un repère géodésique indépendant des distorsions liées aux levés topographiques, nous avons obtenu un géoréférencement extrêmement exact de l'atlas de Verniquet. De plus, nous avons vérifié une des hypothèses de base de notre approche : l'approximation du système de projection ancien par le système actuel Lambert I NTF. À cette occasion, nous avons noté une erreur probable sur le placement de la perpendiculaire à la méridienne de France. Cette erreur n'est cependant pas étonnante étant donné l'approximation initiale de sa position. Enfin, nous avons exploré les possibilités offertes par le géoréférencement pour analyser la qualité géométrique du contenu d'un plan ancien, ce qui nous a permis de fournir une analyse critique de l'atlas de Verniquet.

Avec une erreur d'estimation moyenne de 1.1 mètre sur les 72 planches de l'atlas et une validation de l'approximation de son système de projection, l'approche proposée semble adéquate pour nos plans dès lors qu'ils reposent sur un canevas géodésique de qualité.

Nous avons appliqué la même approche au plan de Maire, à l'atlas de Jacoubet et à l'atlas municipal. Si ces trois plans présentent des carroyages, celui du plan de Maire s'est révélé inutilisable car trop imprécis. En effet, il s'agit d'un carroyage tracé a posteriori sur les feuilles du plan, et non pas un guide géodésique. Un géoréférencement de ses feuilles en s'appuyant sur les rues et les bâtiments s'est révélé également infructueux. Trop imprécis et généralisé, le contenu du plan aboutit à des erreurs de géoréférencement de plus de 10 mètres dans le centre de Paris. Ces erreurs provoquent alors une rupture entre planches, le rendant inutilisable. Pour ces différentes raisons, le plan de Maire a été laissé de côté pour le reste de cette thèse. L'atlas de Jacoubet et l'atlas municipal sont présentés, géoréférencés et assemblés, en annexes D et E

3.6 Vectorisation des plans

Une fois nos plans géoréférencés, nous devons en extraire des données vectorielles qui nous serviront à constituer une base de données géohistorique par plan sur la ville de Paris. Cette vectorisation a été effectuée de façon manuelle. En effet, si une approche automatique avait été envisageable pour un seul plan, leur hétérogénéité ne permet pas d'utiliser cette même approche de manière généralisée. La vectorisation manuelle étant une tâche coûteuse en temps, nous nous sommes limités à la vectorisation d'un seul thème cartographique : les rues de Paris. Nous présentons

3.6.2 Résultats de la vectorisation

Les résultats de la vectorisation manuelle des réseaux viaires des atlas de Verniquet, de Jacoubet et de l'atlas municipal sont présentés dans la figure 3.41. Est illustré également le réseau viaire de l'atlas Vasserot issu des données produites dans le cadre du projet ALPAGE. Chaque réseau viaire est présenté superposé au réseau actuel des rues de Paris³⁶ afin de visualiser l'agrandissement et la densification du réseau au fil du temps. En outre, chaque réseau vectorisé est illustré en grande taille en annexe F.

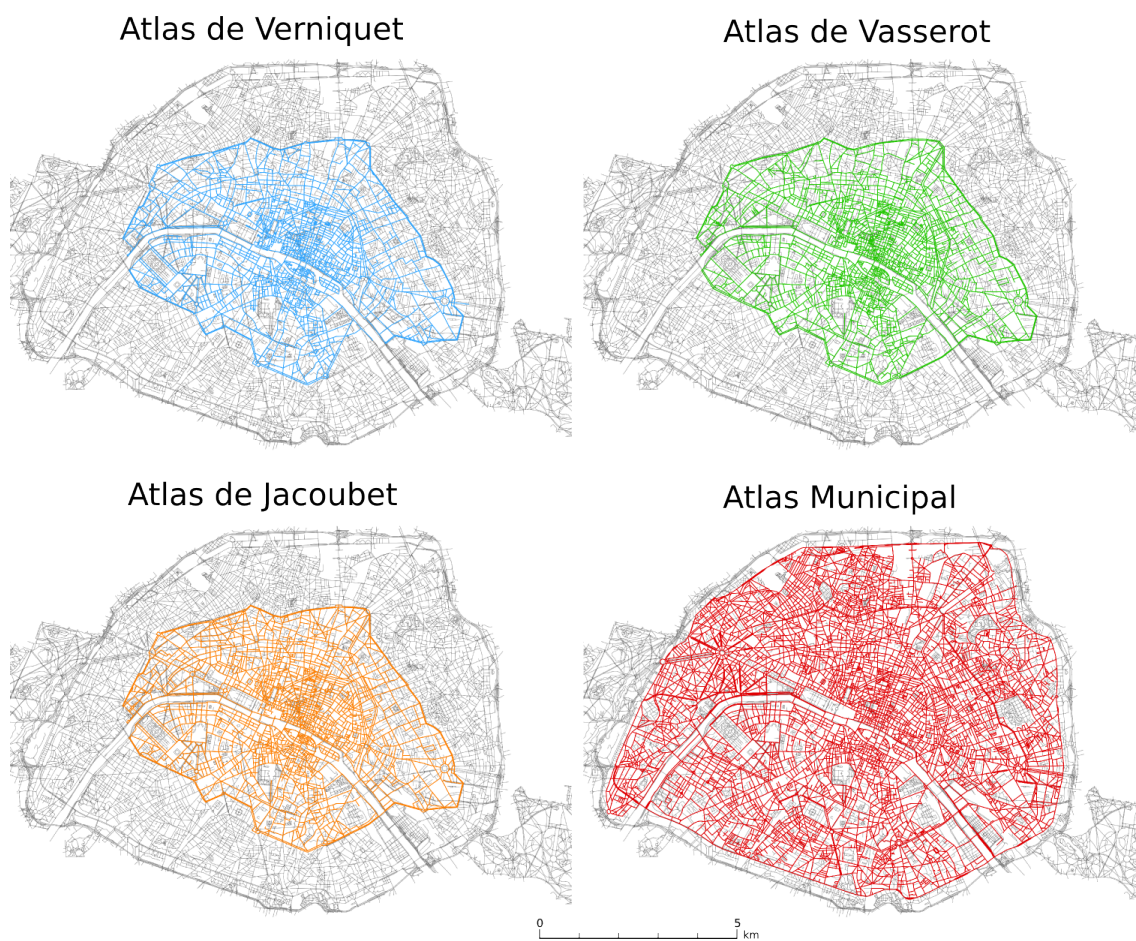


FIGURE 3.41 – Réseaux viaires vectorisés. Le tracé blanc indique l'état du réseau aujourd'hui.

36. Issu de la BDTOPO.

3.7 Conclusion

L'objectif de ce chapitre était de géoréférencer des plans topographiques anciens de Paris dont le système de projection est inconnu, tout en minimisant les erreurs liées aux distorsions géométriques inévitables de ce type de source afin de produire des données vectorielles localisées géographiquement. Une fois produites, ces données constituent le cas d'application du modèle de construction de données géohistoriques proposé dans cette thèse.

Pour atteindre cet objectif, nous proposons une approche de géoréférencement adaptée aux plans anciens de Paris. En particulier, nous montrons qu'il est possible de reconstituer le système de projection de ces plans, ce qui permet alors d'obtenir un géoréférencement de grande qualité et offre la possibilité d'étudier la qualité géométrique interne d'un plan.

L'approche proposée est validée sur l'atlas de Verniquet puis appliquée aux autres plans. Finalement, elle nous permet de géoréférencer tous les plans à l'exception de celui de Maire, trop imprécis.

Au delà du géoréférencement, nous utilisons également cette approche pour étudier en détail la qualité de différents thèmes cartographiques de l'atlas de Verniquet. Cette étude nous permet d'établir un premier modèle des distorsions de deux thèmes de l'atlas : les bâtiments et les rues, mais aussi de pointer la grande qualité de cet atlas et la possibilité d'en faire un référentiel géographique pour l'espace parisien ancien.

Une fois les plans géoréférencés, nous construisons le réseau viaire des rues de Paris sur chacun d'entre eux. Finalement, nous obtenons des réseaux de rues complets, topologiques, sous forme de données vectorielles pour tous les plans à l'exception de celui de Maire. Ces données, groupées par atlas, ne sont pas encore situées dans le temps ni liées entre-elles : elles sont seulement spatiales.

Chapitre 4

Modèle de données spatial et temporel

Objectifs :

- Proposer un modèle de base de données spatiale et temporelle s'appuyant sur les standard de l'information géographique et permettant de structurer les sources géohistoriques et les données vectorielles qui en sont extraites.
- Déterminer la localisation temporelle des sources et des données et en proposer une représentation formelle prenant en compte les incertitudes sur cette localisation.
- Proposer une méthode de tri temporel adaptée aux particularités des données manipulées.

Sommaire

4.1	Modélisation de la base de données spatiale et temporelle . . .	177
4.1.1	La source, l'observation et l'entité géohistorique	177
4.1.2	Modélisation d'une source géohistorique	180
4.1.3	Schéma général de la base de données spatiale et temporelle . . .	183
4.2	Représentation du temps dans les bases de données temporelles	188
4.3	Représentation floue du temps	189
4.3.1	Théorie des sous-ensembles flous	189
4.3.2	Définitions et notations sur le temps flou	194
4.3.3	Modélisation des temporalités valides des sources et observations géohistoriques	195
4.4	Tri temporel	198
4.4.1	Rangement de nombre flous	200
4.4.2	Choix des méthodes de tri temporel et d'antécédence adaptées .	204
4.5	Conclusion	208

Ce chapitre construit une première structure de données spatiale et temporelle visant à organiser et stocker les objets vectorisés dans le chapitre 3 à partir des plans anciens de façon à les rendre utilisables dans un processus de reconstruction des transformations entre eux. Nous suivons finalement la 'montée en niveau' dans l'expressivité spatio-temporelle que nous avons abordée au début du premier chapitre.

Nous introduisons ici un schéma de base de données dans lequel la dépendance forte des objets vectoriels à la carte de laquelle ils sont extraits est conservée. Ainsi, nous adoptons une approche prudente des informations que l'on manipule : elles ne sont pas directement issues de l'observation du monde réel mais extraites d'une première étape d'observation et de restitution cartographique avec toutes les erreurs, filtrages et distorsions de la réalité que cela implique. Nous proposons d'asseoir l'implémentation de cette base de données sur les standards de l'information géographique.

Nous proposons ensuite un façon de modéliser des temporalités incertaines de façon à pouvoir localiser les objets vectorisés et leurs sources cartographiques dans le temps.

Enfin, nous choisissons deux méthodes visant à ordonner et comparer ces temporalités incertaines.

4.1 Modélisation de la base de données spatiale et temporelle

Dans cette section, nous présentons le modèle de données spatial et temporel que nous avons adopté pour stocker les informations vectorisées à partir des plans anciens. Celui-ci s'appuie sur un modèle plus général des notions de "source géohistorique" et "d'observation géohistorique" qu'il spécialise au cas des sources cartographiques historiques et de leur représentation sous forme de données vectorielles. Enfin, nous décrivons comment ce modèle s'intègre aux standards de l'information géographique et les étend afin de manipuler des temporalités incertaines.

4.1.1 La source, l'observation et l'entité géohistorique

Dans cette section, nous présentons un modèle de données simple permettant de représenter les données géohistoriques issues des plans topographiques tout en conservant deux caractéristiques qui leur sont spécifiques : leur forte dépendance à la source, à conserver, et leur statut de *traces*.

Un double biais d'observation

L'OGC¹ fournit un modèle général des opérations de modélisation et de saisie nécessaires à la représentation des phénomènes du monde réel au sein d'une base de données géographique vectorielle. Ce modèle repose sur l'hypothèse que les données sont directement produites par l'observation² des phénomènes du monde réel. Or, dans notre cas, les phénomènes du monde que nous souhaitons étudier ne sont

1. Open Geospatial Consortium : www.opengeospatial.org

2. La notion d'observation étant définie par l'OGC dans le standard *Observations and Measurements* (Consortium, 2003b) comme l'acte de mesurer une propriété d'un phénomène du monde réel.

plus directement accessibles et ne peuvent être appréhendés qu’au travers de traces au sein de sources géohistoriques. Elles-mêmes fournissent une représentation des phénomènes du monde réel issue d’une observation pendant une période donnée. Elles sont donc produites par un processus d’acquisition de données comparable à celui décrit par le modèle général de l’OGC. Leur vectorisation s’apparente donc à une seconde observation succédant à la première et conduisant à une représentation différée et indirecte des phénomènes du monde réel vus à travers le filtre de la source géohistorique. La figure 4.1 adapte le modèle de référence de l’OGC pour modéliser cet enchaînement pour le cas précis des plans de Paris. Les structures urbaines de la ville de Paris sont retranscrites au sein d’un plan de ville à la suite de levés topographiques. Leur représentation cartographique est organisée en thèmes de légende (rues, bâtiments, etc.). Le plan résultant de ces opérations est le fruit d’une observation effectuée sur une période plus ou moins longue. Il est donc représentatif des structures de la ville pendant cette période uniquement, dans le meilleur des cas.

La vectorisation du plan s’apparente à une seconde observation qui retranscrit, au sein d’une base de données géographique vectorielle, les traces des structures urbaines de Paris qui apparaissent sur le plan. Ces traces sont stockées sous la forme d’objets vectoriels, appelés *features* dans le modèle de l’OGC et catégorisés par *feature types*. Cette base de données géographique vise alors à représenter les structures urbaines pour la même période que le plan. Cependant, elles ne saisissent que les traces de ces structures, inévitablement déformées et abstraites au moment de la première observation. En outre, elle se rapproche d’un *snapshot* au sens des bases de données spatio-temporelles (voir le chapitre 5), mais s’en détache par le fait qu’elle décrit l’espace non plus à un instant donné mais pendant une certaine durée.

L’entité géohistorique et sa trace

Nous souhaitons retracer les transformations des entités géohistoriques qui peuplent l’espace urbain. Nous avons vu qu’il était impossible d’observer directement ces entités. Leurs traces sont en revanche accessibles à travers différentes sources géohistoriques, comme par exemple des plans de ville. Il est alors possible de recomposer les transformations des traces des entités géohistoriques au sein d’une succession de sources.

Dans cette thèse, nous formulons l’hypothèse que ces traces capturent l’état des entités qu’elles représentent lors de la création d’un plan. Reconstituer les transformations entre traces reviendrait alors à reconstituer celles des entités géohistoriques sous-jacentes. Plus le nombre de plans et donc d’états différents des entités sont disponibles, plus l’approximation de leurs véritables transformations se rapprochera de la réalité. Disposer d’un grand nombre de plans successifs pour une période donnée revient donc à s’approcher d’une cartographie dynamique des entités géohistoriques et de leurs transformations.

La source géohistorique

Étudier les traces des entités géohistoriques suppose de disposer de sources d’informations au sein desquelles les retrouver. Dans notre cas, ces sources géohistoriques sont des plans topographiques de Paris. Elles sont le fruit d’un premier processus d’observation des structures urbaines et le support de la vectorisation des traces géo-

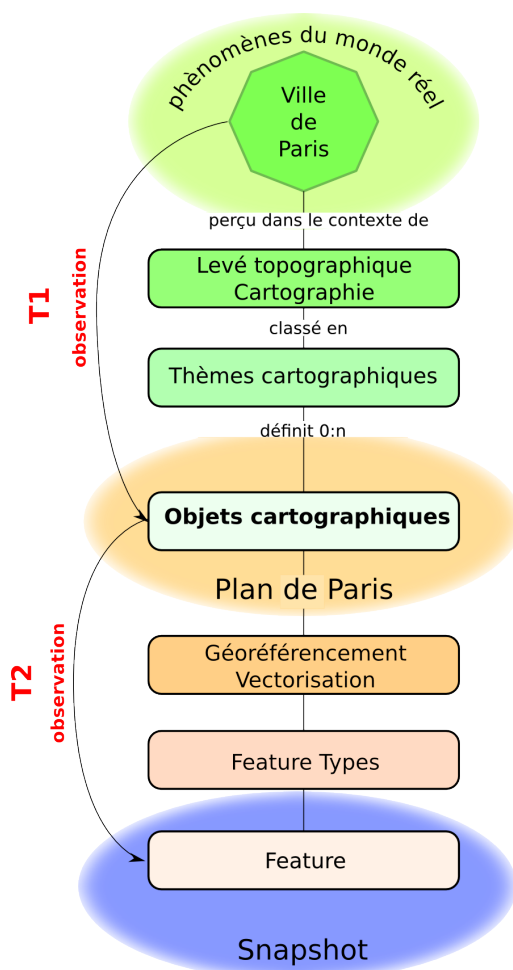


FIGURE 4.1 – Adaptation du modèle de référence de l’OGC pour les sources et observations géohistoriques.

historiques qu’elles contiennent. En outre, les traces des structures urbaines qu’elles contiennent ne sont représentatives de leurs états réels que pour une période à déterminer. Chacune ne peut être considérée comme une source d’informations fiable que pour cette période. Chaque source de donnée géohistorique permet de créer un *snapshot* contenant une représentation des traces.

L’observation géohistorique

Dans cette thèse, nous utilisons le terme d’observation géohistorique pour désigner la représentation des traces des entités géohistoriques contenues dans les sources géohistoriques au sein d’un *snapshot*. Les observations géohistoriques sont donc le produit des deux processus d’observation illustrés en figure 4.1.

La figure 4.2 présente le schéma général décrivant les différentes notions abordées dans cette section.

Généalogie des données géohistoriques

La reconstitution des transformations des entités géohistoriques à partir de leurs

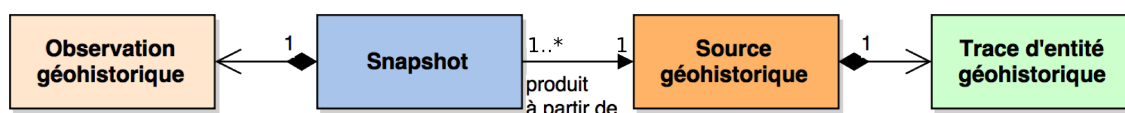


FIGURE 4.2 – Des traces aux observations géohistoriques

traces dans des cartes demande de considérer ces traces avec prudence. Leur fiabilité est liée à celle de la source géohistorique qui les représente. Or, nous avons vu que ces sources étaient imparfaites. Le risque est alors d’aboutir à des reconstructions fausses, et donc inutilisables pour tout travail ultérieur. Pour pallier ce problème, une première possibilité serait de vérifier, par l’intermédiaire d’experts historiens, chaque information avant de l’intégrer dans le *snapshot*. Une telle solution est envisageable pour des *snapshots* de petite taille, mais devient difficile dès lors le nombre des observations augmente, ou que les mises à jour sont fréquentes. Ce modèle prévaut pourtant dans nombre des SIG historiques existants évoqués dans le chapitre 1. La seconde possibilité consiste à intégrer les observations sans expertise, tout en adoptant une stratégie prudente permettant en particulier de conserver un lien fort entre l’observation et la source géohistorique qui a permis sa création. Dans les bases de données géographiques, cette préoccupation est également présente au travers de ce que l’on nomme *généalogie des données* (Bucher, 2005), décrite en particulier dans la norme ISO 19115 (ISO, 2014). Cette norme propose un schéma conceptuel permettant, pour un ensemble de données, de conserver une trace des processus ayant permis sa création. Les sources historiques sous forme d’archives étant au coeur de tout travail d’histoire, il est crucial de pouvoir, à tout moment et pour toute observation, retrouver sa source. C’est pourquoi le schéma que nous proposons en figure 4.2 conserve une dépendance forte entre observations et sources.

4.1.2 Modélisation d’une source géohistorique

Conserver un lien entre observations et sources géohistoriques implique de disposer d’un modèle de représentation de ces sources. Nos uniques sources géohistoriques sur l’espace parisien sont les atlas et plans topographiques anciens de la ville. Les observations correspondent quant à elles à des objets géographiques vectoriels, ou *features*, stockés dans une base de données. La figure 4.3 l’application du schéma présenté en figure 4.2 au cas de la vectorisation de plans. Chacun de ces objets est lié au *snapshot* vecteur auquel il appartient, qui lui-même pointe vers une description de la source cartographique utilisée pour la vectorisation. Cette description suit le schéma conceptuel présenté dans la figure 4.4.

Source géohistorique

Dans notre schéma, une source géohistorique est une classe abstraite correspond à la description d’un document. Il existe des standard de métadonnées pour les sources documentaires, dont le *Dublin Core* (Dublin Core Metadata Initiative,

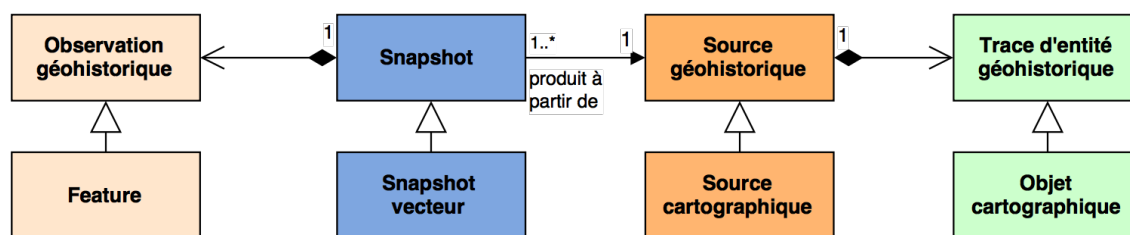


FIGURE 4.3 – Application du schéma observation-source aux plans vectorisés.

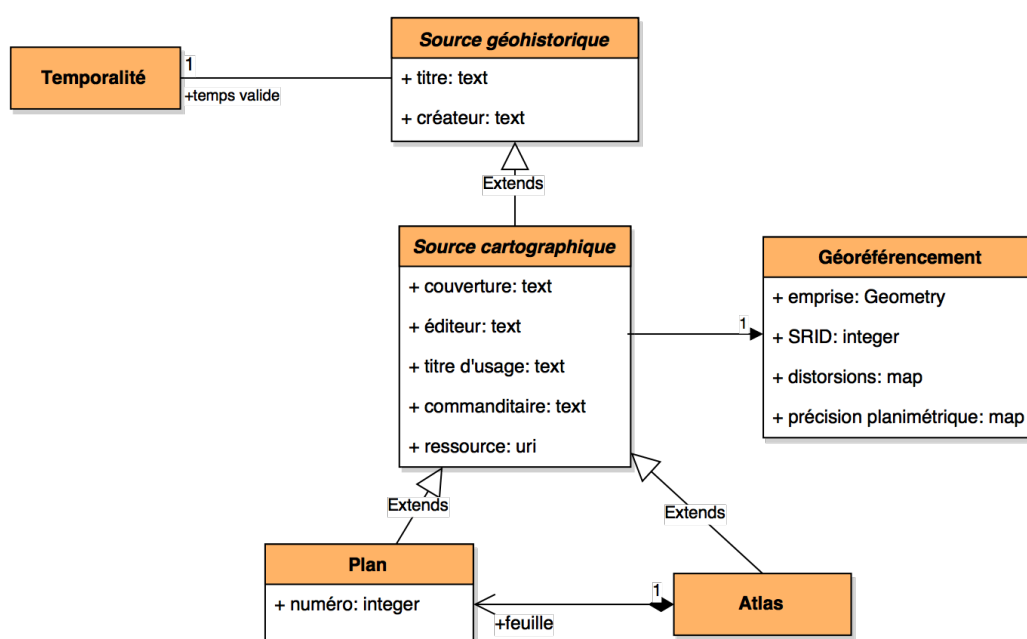


FIGURE 4.4 – Schéma des sources cartographiques.

2012), utilisé pour plusieurs bibliothèques et SIG historiques existants³. Nous avons choisi de réutiliser des éléments de vocabulaire de ce standard pour notre schéma de description des sources cartographiques. Notons qu'une proposition d'extension de la norme de ISO 19115 a été proposée par Éric Grosso (2009) afin de stocker des métadonnées propres aux sources cartographiques anciennes⁴. La classe **source géohistorique** présente les attributs **Titre** et **Créateur** détaillés ci-dessous.

- **Titre** : l'intitulé complet de la source. Pour un plan, celui-ci est généralement situé dans le cartouche. Ce titre peut être relativement long, et n'est pas nécessairement un nom d'usage.

3. Par exemple par la Bibliothèque Nationale de France, le *gazetteer Great Britain Historical GIS* et la *New York Public Library*.

4. Cette extension propose notamment un stockage des informations de qualité du géoréférencement : erreurs résiduelles, précision des points d'amer, processus utilisé. Il serait donc intéressant, dans un second temps, d'intégrer cette extension à notre proposition.

- **Créateur** : l’auteur de la source, s’il est connu. Pour un plan, il peut s’agir du géomètre l’ayant levé, par exemple Edmé Verniquet, ou bien du responsable de sa réalisation, comme Eugène Deschamps pour l’atlas municipal.

Les traces décrites par une source géohistorique doivent être situées dans le temps si l’on veut pouvoir reconstituer leurs transformations à l’aide d’autres traces issues de sources différentes. C’est pourquoi on associe à chaque source géohistorique une temporalité correspondant à la période pour laquelle on peut considérer qu’elle représente une portion de l’espace de façon fiable. On appelle cette temporalité **temps valide** de la source. Par exemple, le *plan archéologique de Paris* réalisé par Adolphe Berty et Robert Lenoir⁵, créé à la fin du XIX^e siècle, reconstitue les maisons et parcelles de la ville médiévale. Le temps valide que nous lui associerions dans notre base serait centré sur la période médiévale. Le temps valide associé à chaque source est donc intrinsèquement lié à celle-ci et doit par conséquent être étudié au cas par cas. La fiabilité de cette temporalité dépend des connaissances dont on dispose sur la source et sur les informations ayant permis sa création. Pour les plans parisiens, nous verrons dans la section 4.3.3 comment il est possible de déterminer ces temps valides.

Source cartographique

Une source cartographique est une spécialisation de la notion de source géohistorique. Elle est également décrite par des attributs en partie issus du modèle Dublin Core, détaillés ici.

- **Couverture** : dans le Dublin Core, la couverture (*coverage*) est une variable qualitative englobant l’étendue spatiale (l’île de la cité, Paris, Union Européenne, etc.) et temporelle (la période 1840-1850 par exemple). Dans notre cas, nous conservons uniquement la composante spatiale, le temps étant géré par une variable quantitative, le temps valide.
- **Éditeur** : Cet attribut désigne la personne ou l’organisation ayant gravé le plan, la plupart du temps différente du créateur. Par exemple, les plans de Paris créés par l’administration de Haussmann sont, pour la plupart, gravés par l’éditeur privé Wuhler.
- **Commanditaire** : en plus de l’auteur et du graveur, il est utile de conserver le commanditaire du plan, en particulier pour pouvoir produire une critique de la source. Déterminer le commanditaire n’est pas trivial car il est rarement écrit sur la source elle-même ou car il n’en existe pas de clairement défini. Enfin, il peut également y avoir plusieurs commanditaires lorsque la création du plan est complexe ou longue. C’est le cas de l’atlas de Verniquet, d’abord effectué au compte du roi Louis XVI, et terminé sous le Directoire.
- **Ressource** : la ressource désigne un moyen d’accès vers le plan scanné et géoréférencé. Pour nous le plan est une métadonnée permettant de conserver la généalogie des features vectorisées. De plus, ce n’est pas la ressource documentaire (le scan par exemple) que nous souhaitons référencer mais sa version géoréférencée car elle correspond au canevas sur lequel les features ont été vectorisées.
- **Titre d’usage** : les plans topographiques anciens ont souvent des noms longs,

5. Dont les feuilles ont été éditées séparément. Par exemple, voir (Berty et Lenoir, 188)

ou très semblables (*plan de la ville de Paris* par exemple). Pour cette raison, ils possèdent également souvent un nom d'usage, plus spécifique et court. C'est par exemple le cas lorsque l'on fait référence à l'*atlas de Verniquet* plutôt qu'à son libellé complet. Ce nom étant simplement une convention d'usage, il peut en exister plusieurs (l'atlas de Verniquet est par exemple également dénommé *Grand Plan de Paris* par les auteurs du XIX^e siècle). Cet attribut n'est pas issu du Dublin Core.

Une information importante à conserver concernant les sources cartographiques concerne leur géoréférencement. En effet, elle permet d'inscrire cette source cartographique dans un système de coordonnées de référence. De plus, cela permet de conserver des métadonnées sur la qualité géométrique de la carte et des features vectorisées sur cette carte. La classe **Géoréférencement** porte ces informations au travers des attributs suivants :

- **Emprise** : l'emprise correspond à la "boîte englobante" de la source géoréférencée. Il s'agit d'une géométrie polygonale décrivant les limites de la source dans l'espace géographique.
- **SRID** : le SRID, *Spatial Reference System Identifier*, est l'identifiant du système de coordonnées de référence dans lequel a été géoréférencé la source. Dans notre cas, il s'agit du code EPSG 27561 qui correspond à la projection Lambert zone I NTF.
- **Distorsions** : la prise en compte des distorsions des plans peut permettre un paramétrage plus fin du processus de reconstitution des transformations entre observations. Ceci nécessite de disposer d'un modèle de distorsions. Nous avons vu au chapitre 3 qu'il était possible d'approximer ce modèle à partir d'un ensemble de points de mesure. Nous proposons ici de le stocker sous la forme d'un raster.
- **Précision planimétrique** : l'étape de géoréférencement fournit une estimation de la précision planimétrique du plan géoréférencé. Cumulée à la précision planimétrique du référentiel géographique utilisé pour le géoréférencement, elle peut également être mise à profit dans le processus de reconstruction des transformations. Cette précision est représentée ici par un raster car elle peut varier sur l'étendue du plan.

Plan et atlas

Compte tenu de leur échelle, certains plans topographiques de la ville ont été découpés en feuilles et rassemblés en atlas de façon à pouvoir être édités. Pour un même plan ainsi décomposé, des graveurs différents peuvent avoir réalisés différentes feuilles. Dans le cas de la carte de Cassini, les commanditaires et les auteurs diffèrent suivant les feuilles. Le géoréférencement, et donc les distorsions et la précision planimétrique peuvent également varier d'une feuille à l'autre. C'est pourquoi nous effectuons une distinction entre l'atlas -ici recueil de parties de plan- et le plan topographique en une seule feuille, qui est homogène pour l'ensemble des propriétés héritées de la notion de source cartographique.

4.1.3 Schéma général de la base de données spatiale et temporelle

Implémentation fondée sur des standards

Afin d'implémenter notre base de données spatiale et temporelle, nous avons choisi de nous appuyer sur le modèle de référence de l'OGC dédié à la modélisation de bases de données géographiques (Consortium, 2003a). Celui-ci impose de regrouper les objets géographiques (*features*) au sein de catégories d'objets de même nature, appelées *feature types*. Chaque *feature types* est décrit par un ensemble de propriétés représentées par des attributs de natures variées (ISO, 2005). Ces attributs permettent de décrire la forme et la localisation des *features* sous la forme de géométries vectorielles (*SpatialAttributeType*, cf. norme ISO (2003)) et leur temporalité à l'aide d'attributs dédiés à la représentation du temps (*SpatialAttributeType*, cf. norme ISO (2002))).

Ces standards sont, pour la partie spatiale, suffisants pour représenter les observations géohistoriques. Toutefois, le standard ISO 19108 ne permet pas de modéliser des temporalités incertaines. C'est pourquoi, dans la suite de ce chapitre, nous choisirons une façon de modéliser de telles temporalités. La figure 4.5 illustre notre schéma de base de données spatiale et temporelle. Nous avons provisoirement nommé le type de l'attribut temporel "temporalité", mais il sera précisé dans la suite de ce chapitre.

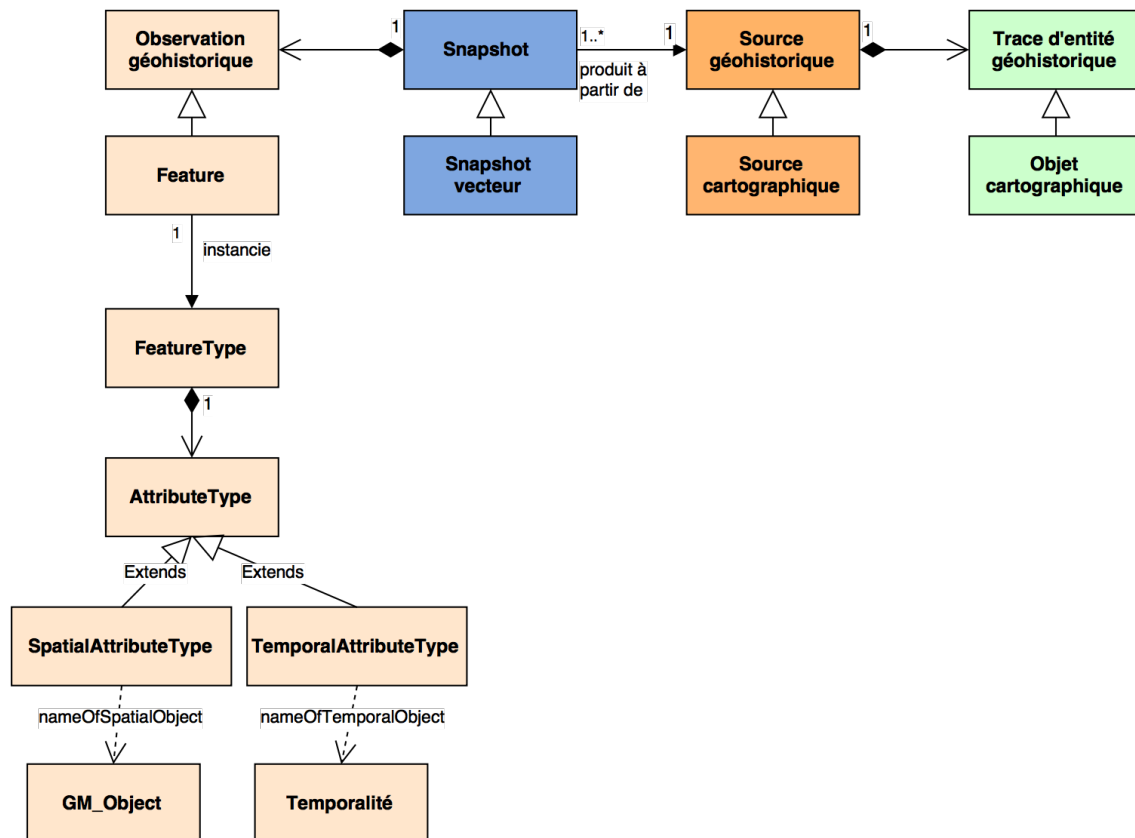


FIGURE 4.5 – Appui sur les standards de l'information géographique.

Instanciation à partir de plans

Notre base est instanciée par vectorisation des sources cartographiques anciennes. Celles-ci présentent des traces d'entités géohistoriques sous forme d'objets cartogra-

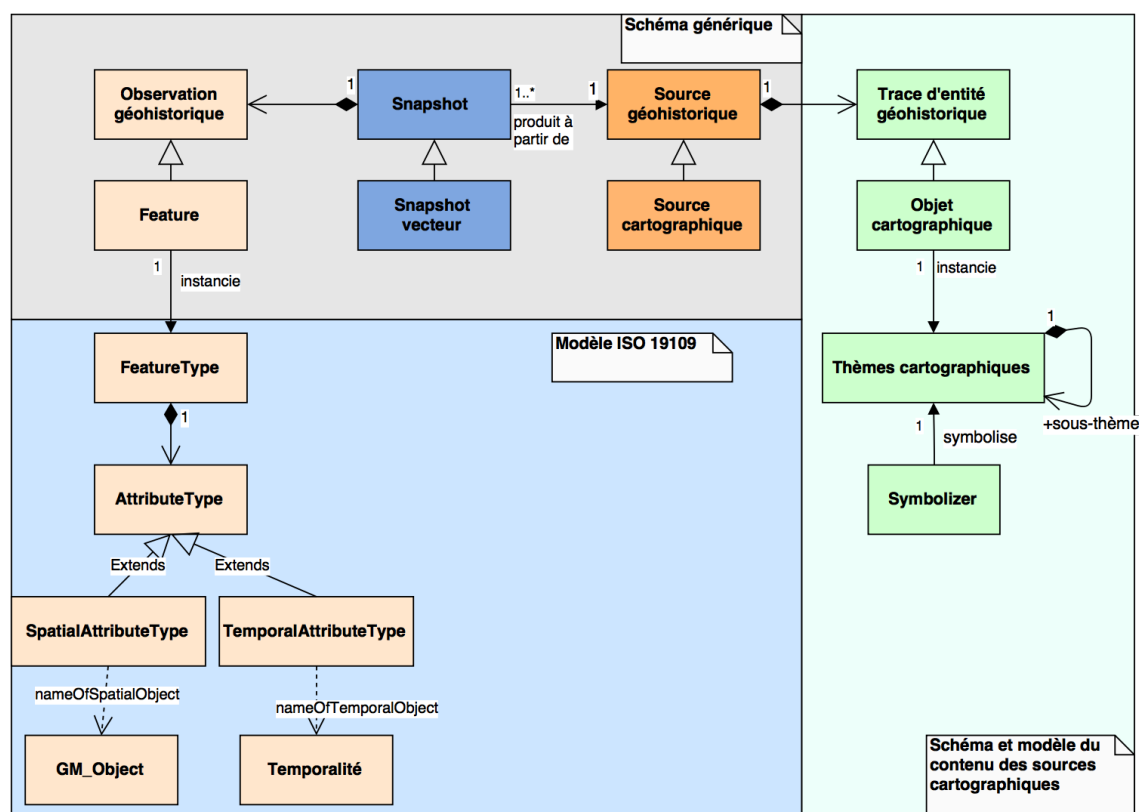


FIGURE 4.6 – Représentation des traces d’entités géohistoriques par des objets cartographiques.

phiques. Ceux-ci sont organisés selon une légende composée de thèmes cartographiques associés à diverses symbolisations. La figure 4.6 complète la figure précédente en lui ajoutant un modèle de description simplifié des objets cartographiques. Seule la partie grisée du schéma concerne des classes pouvant être implémentées et instanciées dans notre base de données. La partie bleutée décrit le standard ISO 19109 pour la modélisation d’objets géographiques que nous réutilisons. Enfin, la partie en vert correspond à un modèle de description du contenu des sources cartographiques. Cette partie du modèle demeure abstraite dans la mesure où les traces sur un plan ne sont pas des objets instanciables dans une base de données. Les classes modélisant les traces d’entités géographiques sont ajoutées ici pour faciliter la compréhension du modèle.

Comment les traces contraignent les observations

L’instanciation de notre base de données spatiale et temporelle est contrainte par les propriétés des sources cartographiques qu’elle représente. En particulier, le choix des types des *featuretypes* est contraint par les thèmes cartographiques disponibles dans les sources (cf chapitre 3). De plus, le choix des primitives géométriques utilisées pour décrire chaque *featuretype* dépend également des symboles utilisés. Enfin, la temporalité associée à chaque *features* doit être représentative de celle de l’objet cartographique à partir duquel il a été vectorisé. De même, la temporalité d’un *snapshot* vecteur est intrinsèquement liée à celle de la source cartographique qu’il

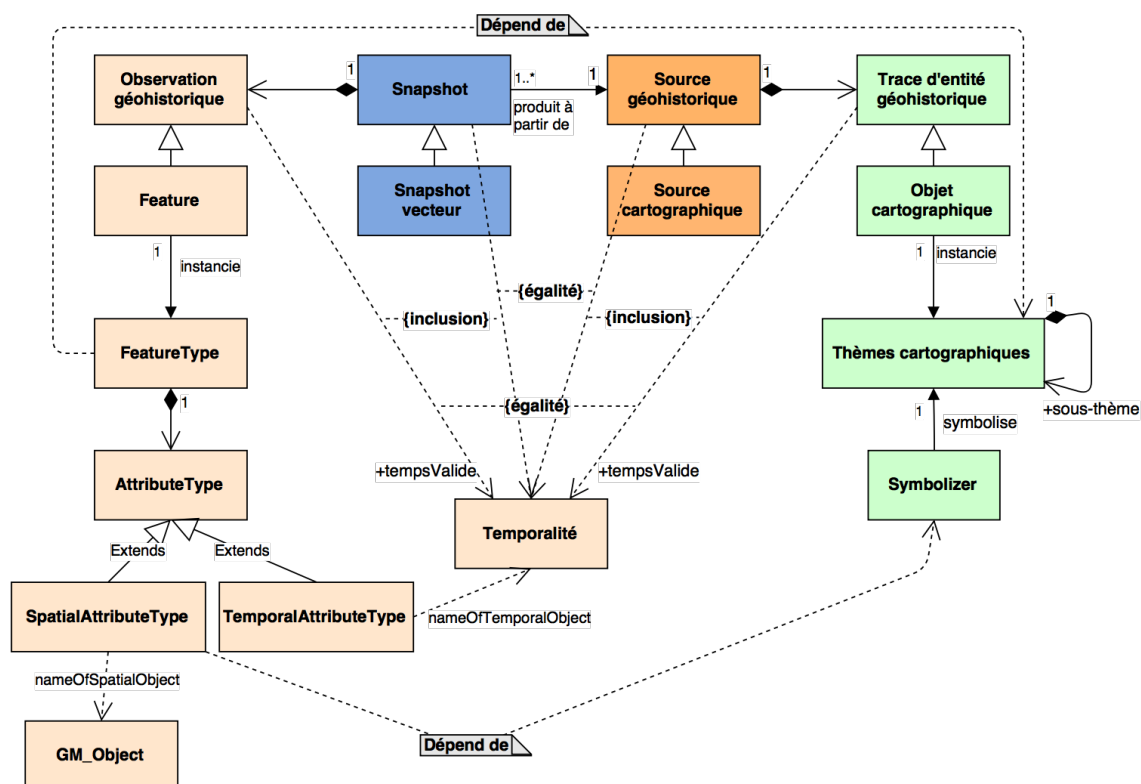


FIGURE 4.7 – Schéma complet d'une base de données spatiale et temporelle.

représente.

Contraintes temporelles observation-source

La trace d'une entité géohistorique est dépendante de la source dans laquelle elle apparaît. Elle correspond à l'observation d'une entité du monde réel pendant le temps valide de la source. Par conséquent, la trace ne peut posséder une temporalité extérieure au temps valide de sa source. Il y a donc une contrainte d'inclusion du temps valide de la trace dans celle de sa source. Cette contrainte se traduit dans notre modèle par une inclusion du temps valide des observations dans celui du *snapshot*. Finalement, le schéma 4.7 résume la totalité de notre modèle, incluant les différentes contraintes énoncées ci-dessus.

Exemple d'instanciation du schéma

Nous présentons en figure 4.8 une exemple d'instanciation de l'observation de la trace d'une portion de rue dans l'atlas de Verniquet modélisée comme un tronçon de rue dans le graphe viaire de Paris. Par souci de simplicité, nous présentons ici le cas où le *snapshot* est défini au niveau de l'atlas complet et non de sa décomposition en feuilles. Le schéma conceptuel de données présenté dans la partie grisée de la figure 4.6 a été implémenté dans le SGBD PostgreSQL muni de son extension spatiale Postgis. Nous avons de plus utilisé une extension permettant de modéliser des temporalités incertaines et leurs relations sous la forme de sous-ensembles flous (voir la section suivante) proposée par Van Daele (2014).

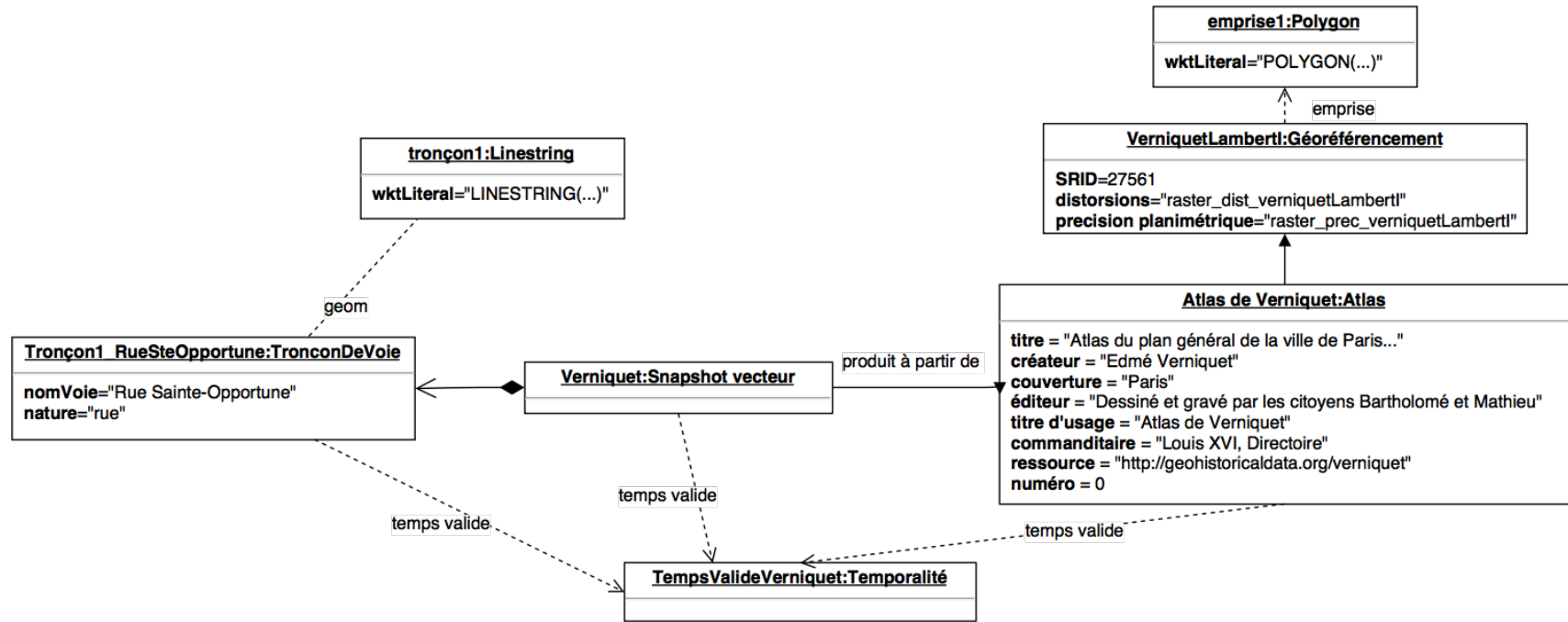


FIGURE 4.8 – Diagramme d’objet illustrant la représentation d’un tronçon de la rue Sainte-Opportune dans la base de données spatiale et temporelle perçue comme une observation d’une trace dans l’atlas de Verniquet.

4.2 Représentation du temps dans les bases de données temporelles

Dans cette section, nous présentons et justifions les choix de modélisation du temps adaptés aux données vectorielles extraites de plans anciens. Pour cela, nous présentons en premier lieu les représentations et modèles de raisonnement sur le temps tels qu'ils existent dans le domaine de l'information géographique, et, plus largement, dans celui des bases de données. Nous présenterons également comment ces modèles peuvent être adaptés à des informations temporelles incertaines, nous permettant d'avoir un mode de représentation et des outils de raisonnement adaptés à des objets géographiques difficiles à situer temporellement.

L'objectif de cette section est de délimiter le plus précisément possible le cadre théorique et méthodologique nous permettant de sélectionner un mode de représentation et de raisonnement sur le temps adapté aux données historiques traitées.

Types de temps

Pour modéliser les transformations d'objets géographiques, il faut ajouter à la représentation de l'espace celle du temps. Le temps est souvent construit par analogie à l'espace, en le ramenant à une seule dimension (Klein *et al.*, 1994). Sous sa représentation la plus simple, il s'agit donc d'une courbe. Dans le domaine des bases de données, la même analogie est à l'origine des modèles de représentation du temps (Frank, 1998; Peuquet, 1994, 2002). Topologiquement, cette courbe peut être fermée, ou non, ce qui détermine deux visions différentes du temps (Frank, 1998). Lorsqu'elle est ouverte, il s'agit d'un temps linéaire, qui se déroule inéluctablement du passé vers le futur. À partir de cette conception linéaire, il est possible de définir des temporalités multiples, d'explorer plusieurs hypothèses (on parle alors de temps ramifié)(Worboys, 1990). Fermée, la courbe détermine un temps cyclique, comme celui des saisons qui se succèdent, ou celui des anciennes civilisations amérindiennes. Il est assez clair que le temps historique, qui décrit les transformations de l'espace, est au moins à notre échelle temporelle linéaire : les événements se déroulent et se succèdent les uns après les autres, sans se répéter. Notons d'ailleurs que la vision cyclique du temps pose problème pour décrire des suites de transformations, puisqu'elle peut aboutir à des situations où le principe de causalité n'est pas respecté. Le temps est associé à deux primitives : l'instant et l'intervalle. Un instant est l'interstice entre le passé et le futur ; dans une vision générale et continue du temps, il est donc de durée nulle. L'intervalle exprime quant à lui une période de temps. L'établissement d'une logique de raisonnement sur le temps dépend du choix d'une primitive temporelle (instant ou intervalle) (Stock, 1998).

En parallèle de ces deux visions conceptuelles, on distingue trois natures du temps dans de nombreux travaux sur les bases de données temporelles (Jensen et Snodgrass, 1996) : les temps *valides*, *de transaction* et *utilisateur*. Le temps valide correspond au temps de la réalité, i.e le moment où un événement se produit. Le moment où l'événement est stocké dans la base de données est un temps de transaction. Le temps utilisateur englobe toutes les autres visions du temps, en particulier le temps perçu, fondamentalement différent du temps physique⁶. La distinction entre temps

6. Voir pour cela les travaux de Klein *et al.* (1994) et Fraisse (1967) en particulier qui distinguent temps

de validité et de transaction est essentielle pour créer des bases de données *versionnées*, c'est à dire qui conservent les changements qui affectent les données. Cette préoccupation a abouti à des modèles de bases bi-temporelles gérant les deux temps à la fois (Worboys, 1998). Les deux lignes de temps, orthogonales, permettent de décrire un plan dans lequel s'organisent les données temporelles. Snodgrass (1992) enfin a proposé deux temps supplémentaires, plus spécifiquement dédiés aux données géographiques : un temps *d'acquisition* (le moment où le phénomène est mesuré) et un temps *cartographique* (le moment où les données sont consultables).

On pourra remarquer que nos plans topographiques peuvent être eux-mêmes interprétés comme des bases de données, mais sans distinction possible entre temps valide et transactionnel. En effet, lorsqu'une entité géographique est décrite sur une carte dont on connaît la période de levé, on ne peut pas savoir, sans informations supplémentaires, si la période pendant laquelle la carte est construite correspond effectivement à une période pendant laquelle cette entité existe. Le temps construction d'une carte peut en effet être plus long que la durée d'existence de l'entité représentée. Il est également possible qu'une carte représente des entités disparues au moment du levé, où bien même n'ayant jamais existé (c'est par exemple le cas des projets de voirie).

La flèche du temps

Qu'il soit conceptualisé comme linéaire ou cyclique, le temps possède deux propriétés fondamentales. Premièrement, il est continu, tout comme la courbe qui le représente, les points qui la compose devenant un ensemble d'instantanés de durée nulle. Il est également infini, vers le passé tout comme vers le futur. Deuxièmement, il est orienté : c'est ce que l'on appelle la *flèche du temps*. Dans le temps, il n'est en effet possible d'aller que du passé vers le futur. Si la première propriété renforce l'analogie temps/espace, la seconde la stoppe. En effet, dans l'espace il est possible de se déplacer dans toutes les directions et de revenir sur ses pas, à l'inverse du temps pour lequel le sens de la marche est imposé.

4.3 Représentation floue du temps

Puisque les informations dont on dispose sur les temps valides des observations et des sources sont incertaines, la chaîne de transformation entre les observations est elle-même incertaine. Il faut donc être capable de modéliser l'incertitude des temps valides des observations et des sources géohistoriques. Pour cela, nous nous appuyons sur la théorie des sous-ensembles flous.

4.3.1 Théorie des sous-ensembles flous

Le concept de sous-ensemble flou (ou plus simplement *ensemble flou*), créé par Zadeh (1965), est la pierre angulaire de la représentation formelle des connaissances imprécises. La théorie des sous-ensembles flous permet, en assouplissant la notion

physique et psychologique. Le premier est celui de la mesure, qui se déroule à un rythme constant, et le second est un temps subjectif dont le rythme varie.

d'ensemble classique, de traduire des situations dans lesquelles un élément n'appartient pas complètement à l'ensemble. L'inconvénient des ensembles classiques est de définir des limites abruptes souvent inadaptées pour traduire des propositions en langage naturel, ou provenant d'instruments de mesure. Par exemple, on peut associer à la proposition "Jean est grand" l'ensemble des tailles supérieurs à 1m90. La classe "grand" possède une limite nette, et Jean n'est pas grand s'il mesure 1m88. Or, il serait plus juste de le considérer comme très proche de la notion de "grande taille". Cette idée de gradation dans l'appartenance à une classe constitue le fondement de la notion d'ensemble flou.

Nous présentons ici la définition, ainsi que quelques propriétés des sous-ensembles flous. Nous spécialiserons cette représentation pour la représentation du temps à l'aide d'ensembles flous. Les éléments présentés ici sont principalement issus de (Bouchon-Meunier, 1995).

Des sous-ensembles classiques aux sous-ensembles flous

Soit un ensemble X quelconque. Un sous-ensemble A de X est défini par une fonction caractéristique χ valant 0 pour tout élément n'appartenant pas à A , et 1 pour les autres. Pour qu'un élément soit imprécis et n'appartient pas totalement à A , il faut que χ puisse prendre des valeurs intermédiaires.

Un sous-ensemble flou \tilde{A} généralise l'ensemble en définissant à la place de la fonction binaire χ une fonction d'appartenance $\mu_{\tilde{A}}$ qui associe à chaque élément x de X un degré d'appartenance compris dans l'intervalle $[0, 1]$. Si $\mu_{\tilde{A}}$ prend seulement les valeurs 0 ou 1, on revient à un ensemble classique. On note $\text{Prop}(A)$ la propriété qu'exprime l'ensemble A , par exemple "Jean est grand".

Il faut ajouter à cette définition quelques caractéristiques des sous-ensembles flous qui sont à la base de l'arithmétique floue :

1. Le **noyau** d'un sous-ensemble flou \tilde{A} est un ensemble $n(\tilde{A})$ contenant tous les éléments de X qui appartiennent totalement à \tilde{A} :

$$n(\tilde{A}) = \{x \in X | \mu_{\tilde{A}}(x) = 1\}$$

2. Le **support** de \tilde{A} désigne l'ensemble des éléments de X appartenant à \tilde{A} :

$$\text{supp}(\tilde{A}) = \{x \in X | \mu_{\tilde{A}}(x) > 0\}$$

3. Enfin, la **hauteur** de \tilde{A} est le plus grand degré d'appartenance pris par un de ses éléments. On dit en particulier que l'ensemble est **normalisé** si sa hauteur est de 1, donc qu'il existe au moins un élément x de X qui appartienne totalement à l'ensemble.

$$h(\tilde{A}) = \sup\{\mu_{\tilde{A}}(x) | x \in X\}$$

La figure 4.9 montre une représentation floue possible de la propriété "Jean est grand", avec ses noyau, support et hauteur. Comme nous n'utilisons pas dans ce mémoire d'ensembles classiques, nous noterons l'ensemble flou \tilde{A} simplement A , et sa fonction d'appartenance μ_A .

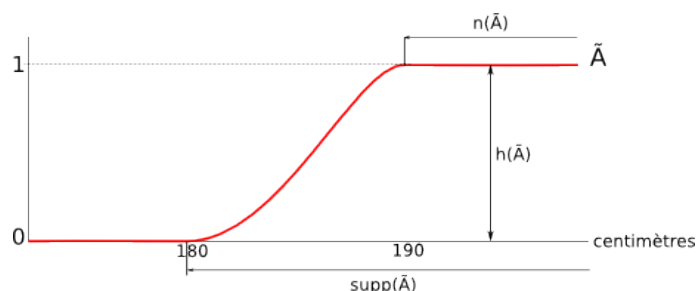


FIGURE 4.9 – Un exemple de sous-ensemble flou associé à la propriété "Jean est grand"

On définit également quelques relations sur les sous-ensembles flous qui nous seront utiles pour la suite du mémoire. Soit deux sous-ensembles flous A et B de X .

1. **Égalité** : Deux sous-ensembles flous sont égaux si leurs fonctions d'appartenance sont égales pour tout élément de X :

$$\forall x \in X \mu_A(x) = \mu_B(x)$$

2. **Inclusion** : L'inclusion $A \subseteq B$ signifie que tout élément x appartenant à A appartient au moins autant à B :

$$\forall x \in X \mu_A(x) \leq \mu_B(x)$$

3. **Union** : L'union $A \cup B$ est un sous-ensemble flou C tel que :

$$\forall x \in X \mu_C(x) = \max(\mu_A(x), \mu_B(x))$$

L'union exprime le fait qu'au moins une des deux propriétés Prop(A) ou Prop(B) est satisfaite ; il s'agit d'un "ou" inclusif. Le degré de satisfaction de Prop($A \cup B$) est donc égal pour tout élément à la propriété la mieux satisfaite.

4. **Intersection** : L'intersection $A \cap B$ est un sous-ensemble flou C tel que :

$$\forall x \in X \mu_C(x) = \min(\mu_A(x), \mu_B(x))$$

La propriété Prop($A \cap B$) est satisfaite lorsque Prop(A) et Prop(B) sont satisfaites simultanément. Le degré de satisfaction de cette nouvelle propriété est donc logiquement celui de la propriété la moins bien satisfaite.

α -coupes d'un sous-ensemble flou

Une façon de manipuler un sous-ensemble flou est de considérer uniquement la partie de ce sous-ensemble satisfaisant un certain degré d'appartenance α . Par exemple, on peut considérer qu'à partir d'un degré de 0.5, Jean est considéré grand. Ce paramètre $\alpha \in [0, 1]$ définit, pour un ensemble flou A , un sous-ensemble A_α nommé α -coupe de A . Ce sous-ensemble est un ensemble classique, défini par sa fonction caractéristique μ_{A_α} :

$$\mu_{A_\alpha}(x) = \{x \in X | \mu_A(x) \geq \alpha\}$$

Sous-ensemble flous convexes et quantités floues

Un sous-ensemble flou A de X est dit **convexe** lorsque pour toute paire d'éléments a, b de X et pour tout réel $\lambda \in [0, 1]$, sa fonction d'appartenance μ_A respecte la contrainte :

$$\mu_A(\lambda a + (1 - \lambda)b) \geq \min(\mu_A(a), \mu_A(b))$$

En découlent deux propriétés essentielles :

1. Le sous-ensemble flou A de X est convexe si toutes ses α -coupes sont également des sous-ensembles convexes de X .
2. L'intersection de deux sous-ensembles flous convexes est convexe.

Lorsque le sous-ensemble flou est normalisé et défini sur l'ensemble des réels, on parle de **quantité floue**. C'est habituellement le cas des mesures imprécises : distance mal estimée, poids imprécis, durée mal connue, etc. De la même façon que sur \mathbb{R} existent des nombres et des intervalles de valeurs, leurs équivalents flous sont nommés **nombres flous** et **intervalles flous**. La figure 4.10a illustre les cas simples d'un nombre flou (à gauche) et d'un intervalle flou (à droite).

Quantités floues Left-Right

Dans les cas les plus simples, les nombres ou intervalles flous à considérer ne sont rien de plus que des ensembles classiques dont les bornes sont définies non plus par une valeur unique mais par deux fonctions, l'une à gauche et l'autre à droite et notées usuellement L et R . Elles sont toutes deux définies sur \mathbb{R}^+ , à valeur dans $[0, 1]$, symétriques, respectivement non décroissante et non croissante, et respectent $L(0) = R(0) = 1$. Le plus souvent, les fonctions choisies sont de types exponentielles ou linéaires qui permettent aisément de créer des fonctions triangulaires, trapézoïdales et gaussiennes (Masson, 2005).

On désigne de façon usuelle ces quantités floues par un triplet $(a, b, c)_{LR}$ pour un nombre flou et $(a, b, c, d)_{LR}$ où $[a, d]$ est le support de l'ensemble flou et $[b, c]$ son noyau. Leurs fonctions d'appartenance, illustrées dans les figures 4.10a et 4.10b, se définissent par :

Pour un intervalle flou :

$$\mu_A(x) = \begin{cases} L\left(\frac{x-a}{b-a}\right) & \text{si } x \in [a, b] \\ 1 & \text{si } x \in [b, c] \\ R\left(\frac{d-x}{d-c}\right) & \text{si } x \in [c, d] \end{cases} \quad (4.1)$$

Pour un nombre flou :

$$\mu_A(x) = \begin{cases} L\left(\frac{x-a}{b-a}\right) & \text{si } x \in [a, b] \\ R\left(\frac{c-x}{c-b}\right) & \text{si } x \in [b, c] \end{cases} \quad (4.2)$$

Cette représentation présente l'avantage d'être facilement représentable informatiquement, tout en facilitant les principales opérations arithmétiques sur les sous-ensembles flous.

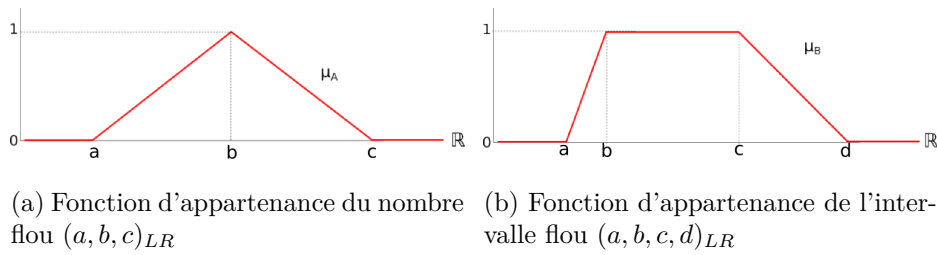


FIGURE 4.10 – Représentation LR d'intervalles et nombres flous.

Principe d'extension de Zadeh et maximum d'un sous-ensemble flou

Le principe d'extension de Zadeh est un principe fondamental de la théorie des sous-ensembles flous car il permet de redéfinir dans le monde flou des relations existantes entre deux ensembles non flous. Si l'on considère deux ensembles X et Y , le principe d'extension définit une fonction Φ qui, à un sous-ensemble A de X , lui associe une image B qui est un sous ensemble de Y . Ce principe général est au fondement de toutes les relations entre sous-ensembles flous, et en particulier l'arithmétique floue. Le principe d'extension est défini comme suit :

$$\mu_B(y) = \sup\{\min(\mu_\Phi(x, y), \mu_A(x)|x \in X\} \tag{4.3}$$

Le principe d'extension de Zadeh permet notamment de définir le sous-ensemble maximum $\widetilde{\max}$, utilisé notamment dans les méthodes de tri de nombres flous que l'on utilise dans ce mémoire. Soient $\{A_1, A_2, \dots, A_n\}$ un ensemble de sous-ensembles flous de X . La fonction d'appartenance $\mu_{\widetilde{\max}(A_1, A_2, \dots, A_n)}$ est définie par :

$$\forall w, \mu_{\widetilde{\max}(A_1, A_2, \dots, A_n)}(w) = \sup_{x_1, \dots, x_n} \min_{i=1, \dots, n} \mu_{A_i}(x_i) \tag{4.4}$$

$$w = \max(x_1, \dots, x_n)$$

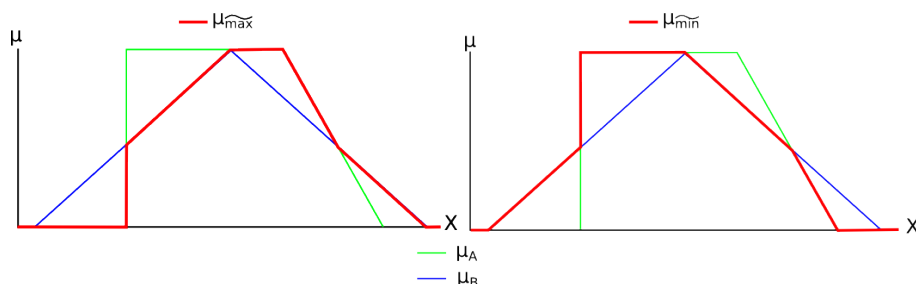
Le maximum d'un grand nombre de sous-ensembles flous étant coûteux à calculer, on peut réécrire ce calcul sous une forme ensembliste plus simple :

$$\widetilde{\max}(A_1, A_2, \dots, A_n) = \bigcup_{i=1}^n \left[\left(\bigcap_{j \neq i} [A_j, +\infty) \right) \cap A_i \right] \tag{4.5}$$

où $[A, \infty)$ est l'ensemble des nombres possiblement supérieurs ou égaux à x , pour tout x appartenant à A . Pour une quantité floue A , $[A, \infty)$ est le sous-ensemble flou défini par $[A, \infty) = \sup_{y \leq x} \mu_A(y)$. Il existe de la même façon $(-\infty, A] = \sup_{y \geq x} \mu_A(y)$.

Dans la suite du document, on notera plus simplement ces deux ensembles respectivement \underline{A} et \overline{A} .

Notons que la relation duale $\widetilde{\min}$ peut se définir de la même façon à partir de \overline{A} . La figure 4.11 donne l'exemple du maximum et du minimum de deux quantités floues. Le maximum et minimum de deux sous-ensembles flous sont particulièrement intéressants lorsqu'il s'agit de comparer ces deux sous-ensembles. En effet, ils renseignent sur la dominance d'un sous-ensemble sur l'autre. Nous verrons que cette relation de dominance est largement utilisée dans le cadre de la comparaison de sous-ensembles flous.

FIGURE 4.11 – \widetilde{max} et \widetilde{min} de deux quantités floues.

4.3.2 Définitions et notations sur le temps flou

Nous introduisons ici quelques définitions et notations utiles pour les prochaines sections et le reste du mémoire.

4.3.2.1 Domaine temporel

Considérant une représentation linéaire du temps, le domaine temporel est noté :

$$\langle T, \leq \rangle \quad (4.6)$$

Le domaine temporel est un ensemble dénombrable d'instants $\{t_1, \dots, t_n\}$ muni de la relation d'ordre \leq équivalente à la relation \leq sur \mathbb{R} .

Ainsi, $t_i \leq t_j$ signifie que l'instant t_i **précède** t_j .

4.3.2.2 Domaine temporel flou

Lorsqu'il y a une incertitude ou une imprécision sur une information temporelle, on peut la représenter au moyen d'une quantité floue, qu'il s'agisse d'un instant (par exemple 19h30 à une granularité d'une minute) ou d'une durée (48 heures). Plus exactement :

- Chaque instant t de T peut s'exprimer sous la forme d'un nombre flou.
- Chaque durée $[t_1, t_2]$ de T peut s'exprimer sous la forme d'un intervalle flou.
- La propriété $Prop(A)$ associée à un sous-ensemble flou A de T est un **événement**. Par exemple, "Jean prend le train" est un événement associé à l'ensemble flou $A = (19h50, 20h, 20h10)_{LR}$ si son train part "environ à 20 heures".

Relations d'antécédence et d'inclusion

En prenant en compte des informations incertaines, le domaine temporel devient alors :

$$\langle T, \preceq, \subseteq \rangle \quad (4.7)$$

Comme nous l'avons défini, T est un homéomorphisme de \mathbb{R} , et, par abus de langage, nous utilisons le terme de quantité floue pour désigner un nombre ou un intervalle flou de T . On a alors :

1. La relation \preceq est un ordre partiel exprimant l'**antécédence**. Pour A et B deux quantités floues de T , $A \preceq B$ signifie que l'événement associé à A se déroule avant celui associé à B . Si on associe "Jean arrive à destination" à B , on a *a priori* $A \preceq B$. \prec . Si on considère seulement l'ordre strict \prec , il est équivalent à la relation *is before* de l'algèbre de Allen. L'égalité entre deux temporalités floues A, B est notée $A \sim B$
2. \subseteq exprime l'inclusion de deux intervalles temporels flous.

On note finalement \tilde{T} l'ensemble des quantités floues existantes sur T .

4.3.3 Modélisation des temporalités valides des sources et observations géohistoriques

Quelle temporalité choisir ?

Pour pouvoir situer les objets géographiques tracés dans un plan sur la ligne du temps, il faut d'abord être capable d'y placer le plan lui-même. En effet, de la même manière qu'il délimite une portion de l'espace, il ne capte la réalité physique que pendant un laps de temps. Le plan se comporte en quelque sorte comme une prise de vue en pose longue : de la même manière que l'appareil photo accumule et superpose les images des objets de la scène, la carte capture des entités à différents moments⁷.

Cette période de temps est particulièrement intéressante, puisqu'elle correspond à moment où le message cartographique est le plus proche d'une représentation de l'espace réel : il y a une forme de synchronisation entre les deux. Cependant, contrairement à une photographie, cette synchronisation peut avoir lieu plusieurs fois. En effet, il est courant qu'un plan, surtout un grand atlas, soit corrigé et mis à jour plusieurs fois. Chaque mise à jour, généralement concrétisée par de nouveaux levés, correspond à une nouvelle synchronisation. Entre deux mises à jour le plan est dans une phase de conservation, la représentation que fournit le plan s'écarte petit à petit de la forme réelle de la ville, au fur et à mesure que celle-ci change.

La succession de conservations et de mises à jours constitue la ligne de vie d'une source cartographique. La figure 4.12 représente cette ligne de vie dans le cas d'un plan levé puis mis à jour plusieurs fois. Les différentes phases qui forment cette ligne de vie sont illustrés dans la partie supérieure de la figure. Chaque phase de levé topographique, qu'il s'agisse du levé originel ou d'une mise à jour, est suivie d'une phase de gravure. L'indication du type de temps décrit la catégorie de temps de chaque phase lorsque l'on considère la source comme une base de données contenant des représentations d'objets géographiques. La partie inférieure illustre l'évolution de l'écart entre la représentation de l'espace dans la source et celle que l'on peut avoir de l'espace réel. Dans un premier temps, celui du levé, les deux sont très proches. Une fois cette première phase terminée, le plan continue pendant un temps à décrire correctement la ville, celle-ci se transformant relativement lentement. Petit

7. Toutefois, à la différence de la photo, la phase d'acquisition d'un plan topographique fait intervenir des facteurs humains qui déforment le message. Voir le paragraphe 1.2.3.1 sur la communication cartographique différée

à petit, le plan s'écarte de l'espace réel, jusqu'à ce qu'une phase de mise à jour ne le resynchronise avec la réalité. Plus le temps sans mise à jour est important, plus la désynchronisation est importante : au bout d'un certain moment, certaines zones décrites par la carte n'existent plus dans la ville réelle. Dans le cas de cartes historiques, on se trouve souvent dans ce dernier cas et, comme indiqué à l'extrémité droite de la ligne du temps dans la figure 4.12, les données extraites de la carte présentent une différence importante avec la ville d'aujourd'hui. Puisque l'on cherche

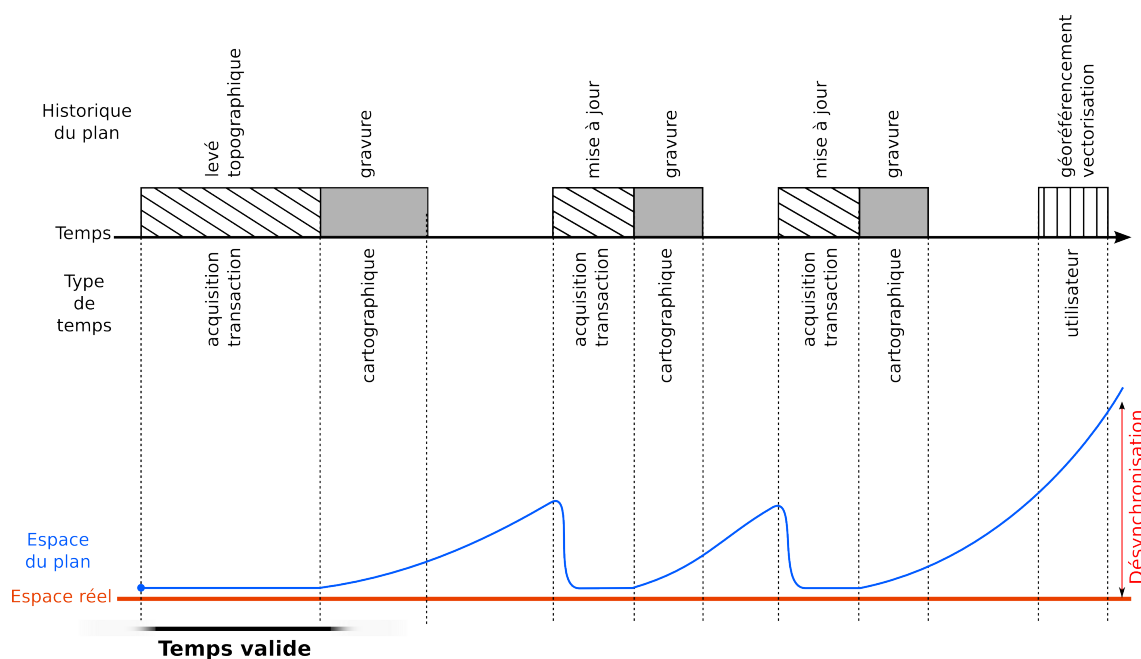


FIGURE 4.12 – Les différentes phases d'existence d'un plan topographique.

à retracer les transformations de l'espace urbain, la période la plus intéressante est celle où le plan représente le plus fidèlement possible la réalité. Le choix se porte donc sur une des phase de levé topographique. Si le plan a subi plusieurs mises à jour, il faut alors considérer que l'on a affaire à des versions successives du plan. Cette suite s'inscrit alors dans un temps de transaction et nécessite donc une base gestion bitemporale des sources géohistoriques. Nous nous limitons cependant ici au seul temps valide, ce qui impose de choisir une des période de levé cartographique comme temps valide. Les plans dont nous disposons n'ont pas subi a priori de mise à jour notable (à l'exception du plan de Maire, que nous ne traitons pas, qui est lui-même une mise à jour d'un plan plus ancien).

Le temps valide de nos plans est donc défini comme leur période de levé topographique.

Notation floue du temps valide d'un plan

On exprime, pour chaque plan p , son temps valide par un intervalle flou trapé-

zoidal $T_p = (a, b, \alpha, \beta)$ dont la fonction d'appartenance μ_{T_p} est définie par :

$$\mu_{T_p}(t) = \begin{cases} 1 - (a - t)/\alpha & \text{si } a - \alpha \leq t \leq a \\ 1 & \text{si } a \leq t \leq b \\ 1 - (t - b)/\beta & \text{si } a \leq t \leq b + \beta \\ 0 & \text{sinon.} \end{cases} \quad (4.8)$$

Le noyau $[a, b]$ correspond à la période de levé topographique, durant laquelle le contenu du plan est considéré comme une représentation fiable de l'espace réel (modulo les inévitables erreurs de levé, etc.). α et β expriment l'incertitude temporelle sur les bornes de cet intervalle résultant :

- de l'imprécision ou de l'incomplétude de l'information temporelle.
- de la tolérance sur la proximité plan/espace réel. β en particulier permet de prendre en compte la période qui suit immédiatement la fin des levés.

Les valeurs de a et b sont propres à chaque plan et sont déterminées d'après les connaissances historiques disponibles. La période de levé est parfois inconnue, soit car le plan est trop ancien, soit car il n'est pas bien documenté. On peut alors s'appuyer sur une analyse historique du plan, comme on l'a fait précédemment pour les planches de l'atlas de Jacoubet, ou utiliser la date de gravure et publication du plan.

La figure 4.13 présente un exemple de temporalisation défini pour un plan dont la phase de levé est représentée par une bande hachurée et celle de gravure en gris. Les valeurs a, b, α et β indiquées doivent être choisies à partir des connaissances historiques sur la source. Par exemple, dans le cas de l'atlas de Verniquet, le temps valide correspondant aux connaissances rassemblées dans le second chapitre nous amène à définir le temps valide $(1783, 1785, 1791, 1799)_{LR}$. Nous avons en effet vu que le tracé du grand atlas s'était étalé de 1785 à 1791, en notant cependant que des premières mesures avaient été effectuées dès 1783. On choisit ici de plus de considérer le plan encore valide jusqu'à la fin de sa gravure, en 1799.

Temps valide des sources utilisées

Le tableau 4.1 présente les temporalités floues assignées aux différentes sources géohistoriques utilisées dans cette thèse. Les bornes de ces intervalles temporels ont été définis à partir des informations collectées au chapitre 2.

Temporalité des observations

Nous avons vu dans le premier chapitre que le contenu des plans ne doit pas être considéré comme un ensemble d'objets géographiques classiques, mais comme des *traces* laissées, à un moment donné, par une entité géohistorique dans une source. Prenons par exemple le cas d'une rue dessinée dans un plan et possédant donc une certaine représentation, une forme et peut être quelques attributs tel qu'un nom ou une date de percement. Sa vectorisation intègre inévitablement les choix de représentation du plan, en particulier ceux de généralisation, mais aussi d'éventuelles erreurs. De plus, le temps valide du plan ne couvre souvent qu'une petite portion de la période où la rue existe. Il est peu prudent, historiquement, de considérer que la forme qu'elle possède sur le plan est correcte pour des périodes où le plan ne

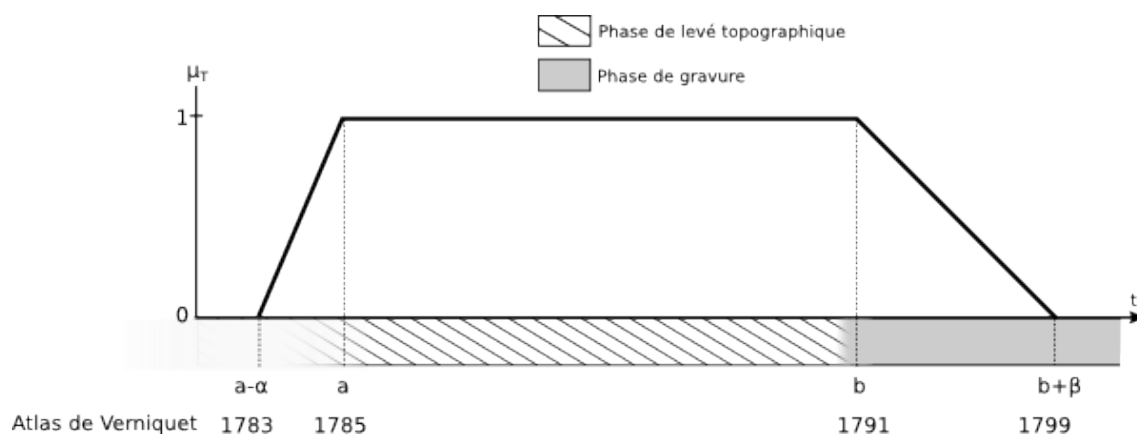


FIGURE 4.13 – Le plan est temporalisé par sous-ensemble flou (a, b, α, β) , choisi pour correspondre à la période de levé topographique durant laquelle les formes urbaines tracées sur le plan sont considérées comme des représentations fiables des structures réelles. Les paramètres α et β sont des valeurs d'incertitude temporelle autour des bornes de validité du plan.

représente plus l'espace réel de manière fiable.

De la même façon qu'une observation est dépendante d'une source, sa temporalité est également dépendante de celle de la source. En effet, s'il s'agit d'une trace d'un objet géographique, celle-ci ne peut exister qu'en même temps que le plan.

La sémantique de cette temporalité diffère cependant de celle de la source. En effet, nous considérons jusqu'ici une période de validité, pendant laquelle l'opération de levé cartographique est effectuée. La temporalité d'une observation doit, quand à elle, exprimer une période pendant laquelle l'objet géographique dont elle est la trace a pu être observé. Il s'agit d'une **possibilité d'observation**.

A priori, l'observation d'un objet géographique est possible pendant toute la durée valide d'un plan. Pour cette raison, la temporalité des observations contenues dans une source est considérée par défaut comme étant celle de cette source. Cependant, il est possible que des connaissances supplémentaires permettent de spécifier cette temporalité. C'est par exemple le cas si l'on sait que la rue observée a été détruite au cours du levé du plan, la possibilité de l'observation étant dès lors nulle.

Il est donc intéressant de permettre aux temps des observations d'être plus restreints que celui de leur plan. Du fait de leur statut de trace, il n'est cependant pas possible que cette temporalité soit plus large que celle de la source. Ainsi, cette relation est exprimée dans le schéma en figure 4.7 par la contrainte d'inclusion entre la temporalité de l'observation et celle de sa source. En effet, la première est nécessairement un sous-ensemble du temps valide de la source, il s'agit donc d'une inclusion entre sous-ensembles flous.

4.4 Tri temporel

Une requête de base lorsque l'on manipule un ensemble d'informations temporelles consiste à les ordonner sur la ligne du temps. Dans notre base de données, il est nécessaire de pouvoir ordonner les sources géohistoriques dans le temps. Dans le

TABLE 4.1 – Temporalités valides des plans utilisés dans la thèse.

Nom d'usage	Temps valide	Représentation graphique
Atlas de Verniquet	(1783, 1785, 1791, 1799)	
Atlas Vasserot	(1807, 1810, 1836, 1854)	
Atlas de Jacoubet	(1825, 1827, 1836, 1837)	
Atlas Municipal	(1887, 1888, 1889)	

cas d'instant, la solution est triviale, mais si elles sont représentées par des intervalles il n'existe plus d'ordre total. Lorsque ces informations sont précises et donc modélisées par des intervalles classiques, la relation *is before* de l'algèbre temporel de Allen (1983) définit ainsi un ordre partiel sur T (Donnelly et Guizzardi, 2012). Si elles sont imprécises, deux solutions s'offrent à nous : soit on modélise l'imprécision directement dans l'information, soit on l'exprime à travers les relations définies dans le raisonnement temporel. Dans le premier cas, une façon de faire consiste à exprimer les connaissances temporelles sous la forme de quantités floues puis de redéfinir l'algèbre temporel dans le cas flou. Dubois *et al.* (2003) en ont ainsi proposé une formalisation possibiliste⁸ permettant de raisonner sur des informations temporelles incertaines et imprécises. Dans le second cas, les intervalles restent classiques, mais ce sont les relations elles-mêmes qui sont rendues floues. Schockaert *et al.* (2008) en particulier a proposé des formalisations de ce type.

Degré d'antécédence entre périodes temporelles floues Ordonner des événements décrits par des quantités floues revient à définir la relation d'antécédence \preceq . Cependant, le simple tri ne suffit généralement pas lorsque l'on traite avec des événements situés dans le temps long, mais il est intéressant de pouvoir quantifier la relation d'antécédence. Autrement dit, il est non seulement utile de savoir si un événement se déroule avant un second, mais également de "combien" il le précède. Ce **degré d'antécédence** ajoute une information à un ordonnancement temporel et peut servir d'aide à la décision, lorsqu'il s'agit de trier par exemple des connais-

8. La théorie des possibilités est une approche de l'incertain élargissant la théorie des probabilités en permettant d'exprimer des incertitudes sur un événement. Elle définit en particulier les notions de possibilité et de nécessité : un événement peut être possible même si la probabilité qu'il se produise est faible (ou inconnue). La nécessité est la mesure duale de la possibilité et exprime quant à elle la certitude que l'événement se produise. Gagner au loto est un événement très peu probable, tout à fait possible mais -en principe- non nécessaire (et parfaitement incertain).

sances archéologiques (De Runz, 2008; Kauppinen *et al.*, 2010).

On note ant la fonction définie sur $\tilde{\mathcal{T}}^2$ mesurant un degré d'antécédence entre deux temps flous. Ce degré d'antécédence est pris dans un intervalle $[0, a_{max}]$, $a_{max} > 0$. Pour deux temps flous A et B pris dans $\tilde{\mathcal{T}}$, on a :

- $ant(A, B) > ant(B, A)$ lorsque A précède B ,
- $ant(A, B) = ant(B, A)$ si A et B sont égaux ou si A précède autant B qu'il lui succède (leurs centres de gravité sont égaux),
- $ant(A, B) < ant(B, A)$ lorsque A succède à B ,

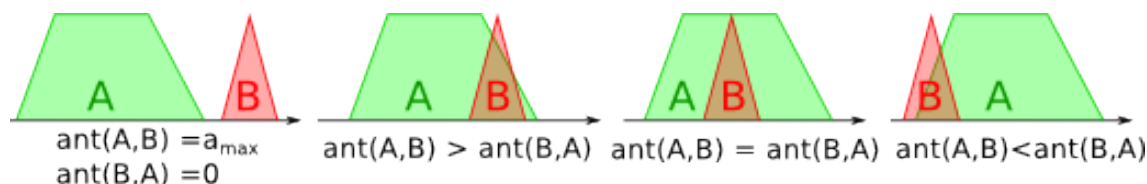


FIGURE 4.14 – Quelques exemples de positions relatives entre temps flous.

Une possibilité pour définir la relation d'antécédence \preceq consiste à utiliser l'une des nombreuses méthodes de tri de nombres flous existantes. Le degré d'antécédence ant peut également être exprimé par certaines de ces méthodes. Nous proposons ici d'explorer les principales mesures existantes pour, une fois la modélisation temporelle de nos données exposée, en choisir une satisfaisante pour notre cas d'étude.

4.4.1 Rangement de nombre flous

Classification des méthodes de tri

Wang et Kerre (2001) classent les méthodes de tri de quantités floues en trois catégories principales :

1. Les méthodes qui projettent les nombres flous sur \mathbb{R} en convertissant chaque nombre flou en un nombre réel, et qui effectuent le tri parmi les nombres réels ainsi calculés.
2. Les méthodes qui comparent chaque nombre flou avec un sous-ensemble flou de référence, puis qui trient les nombres en fonction du résultat de cette comparaison.
3. Enfin, les méthodes qui comparent les sous-ensembles deux à deux et traduisent une relation de dominance. Il faut alors disposer d'une méthode de décision pour trier finalement ces nombres.

Chaque catégorie a des inconvénients. Pour la première, on perd la notion de flou et donc une quantité importante d'information, ce qui est un problème puisque *a priori*, l'imprécision est au centre des informations traitées (Freeling, 1980). Pour la seconde, le problème vient de la signification de l'ensemble de référence : s'il n'a pas un sens aisément compréhensible, on risque d'avoir des résultats contre-intuitifs. Enfin, pour la troisième classe de méthodes, Wang et Ruan (1995) remarquent qu'elles ne sont généralement pas transitives et empêchent de définir un ordre total sur les ensembles considérés. Le rangement de quantités floues est complexe car la préférence d'une alternative imprécise à une autre est une question subjective. Pour cette

raison, il n'existe pas de consensus sur une méthode dans la mesure où elles donnent toutes des résultats allant parfois à l'encontre d'un classement "naturel". Un autre problème vient du fait qu'elles peuvent donner un ordre strict dans des situations complexes pour lesquelles il n'existe pas de choix clair. Pour cette raison, il existe une littérature extrêmement abondante sur le sujet depuis la fin des années 1970. En 2001, Wang et Kerre (2001) comptent plus de 35 propositions majeures, et nous en avons dénombré au moins 11 nouvelles dans les revues spécialisées sur le sujet. Souvent, il s'agit de nouvelles propositions qui visent à corriger certains résultats jugés contre-intuitifs par une méthode antérieure. Quelques auteurs ont proposé des synthèses comparatives de ces différentes méthodes ainsi que des *benchmarks*, en particulier Bortolan et Degani (1985), Chen et Lu (2001a) et Tang *et al.* (2010), dans lesquels le lecteur trouvera une étude approfondie de la plupart des méthodes existantes. Nous ne citerons ici que les méthodes principales issues de ces articles de synthèse, ainsi que les méthodes récentes qu'ils ne considèrent pas. Nous préférons ici les présenter selon les liaisons qu'elles entretiennent (généralisation, correction, équivalence)⁹ plutôt que simplement selon les trois catégories énoncées. Un code couleur indiquera cependant la catégorie à laquelle appartient la méthode, puisque une partie des limites des méthodes sont communes à la classe entière. Les catégories auxquelles appartiennent les méthodes sont cependant indiquées par un code couleur : **violet** pour les méthodes transformant les sous-ensembles en réels, **bleu** pour celle s'appuyant sur un ensemble de référence et enfin **vert** pour les méthodes fondées sur la dominance entre nombres.

Méthodes de tri existantes

Parmi les premières, la proposition de Jain (1977) consiste à comparer chaque quantité floue avec un sous-ensemble maximal de référence correspondant à la notion d'un "nombre aussi grand que possible" pour fournir un indice de rang. Chen (1985) étend par la suite la méthode de Jain pour les quantités floues négatives, et Kim et Park (1990) ajoute une seconde mesure avec cette fois le sous-ensemble minimal. Kerre (1982) reprends l'idée de Jain en calculant la distance de Hamming entre les quantités floues et le sous-ensemble \widetilde{max} . Cependant, comme l'ont montré Baldwin et Guild (1979a), l'ensemble \widetilde{max} est souvent trop peu discriminant et génère des résultats parfois incohérents.

Parallèlement, Baas et Kwakernaak (1977) décrivent une relation de préférence entre deux quantités floues. Baldwin et Guild (1979b) remarque cependant un comportement trop optimiste et propose une relation similaire pouvant être modulée par un "degré de prise de risque". Parmi les méthodes effectuant une transformation des quantités floues en réels et régulièrement reprises, on retrouve notamment Adamo (1980), Yager (1981), Chang (1981), Liou et Wang (1992), Choobineh et Li (1993) et Fortemps et Roubens (1996). Simple, la méthode de Adamo (1980) assigne à chaque quantité floue un indice défini comme la borne supérieure d'une α -coupe

9. De nombreuses méthodes de tri d'ensembles flous ont été proposées, avec l'objectif de produire des rangements proches de l'intuition humaine. Or, aucune n'atteint cet objectif dans tous les cas possibles. De plus, la notion même d'un rangement intuitif est parfois complexe à définir, notamment lorsque les nombres flous sont très proches. En conséquence, de nombreuses méthodes ont été proposées pour corriger les résultats contre-intuitifs de propositions antérieures.

d'un niveau quelconque choisi préalablement. Yager (1981) définit quatre indices comparant pour trois d'entre eux les valeurs de $\mu(x)$ et la valeur de x , considérant que plus x est grand, plus il est important et doit influencer le rang, idée reprise par Chang (1981). Le dernier indice prend quant à lui en compte l'aire entre l'axe des ordonnées et l'axe moyen de la quantité floue. Liou et Wang (1992) considère le cas de quantités floues L-R et propose de comparer les fonctions gauche et droite.

Pour comparer deux quantités floues, Tseng et Klein (1989) introduit en plus des dominances à droite et à gauche la notion d'indifférence qui permet une comparaison plus fine, en particulier lorsque les deux quantités se recouvrent. La même idée se retrouve chez Kołodziejczyk (1986) qui propose trois relations de dominances dont l'une prend également en compte une zone d'indifférence entre les quantités floues, définies par la zone d'intersection entre elles. Chen et Lu (2001b) ont proposé de mesurer la dominance d'une quantité sur l'autre en s'appuyant sur l'écart entre les bornes de leurs α -coupes. Dubois et Prade (1983) ont décrit une approche possibiliste fondée sur quatre indices exprimant des possibilités et nécessités de dominance entre deux quantités floues (à droite et à gauche). Le choix d'une priorité dans les indices pour déterminer un ordre final est laissé libre, et Bortolan et Degani (1985) ont montré qu'ils sont, dans le pire des cas, trop peu discriminants. D'autres théories des connaissances imparfaites ont été utilisées, celle des probabilités par Lee et Li (1988) dont la mesure a été corrigée par Cheng (1998), ou encore la théorie des fonctions de croyance par Sevastianov (2007). Dans le but de comparer des sous-ensembles flous représentant des intervalles de temps, De Runz (2008) propose une adaptation de Kerre, fondée sur la relation de dominances entre quantités floue. Il introduit ainsi un indice d'antécédence entrs sous-ensembles temporels flous. Ces méthodes, outre le fait de s'appuyer sur des bases mathématiques solides, présentent l'avantage de permettre une comparaison fine et souple de deux quantités floues, mais aussi de fournir des mesures sur la qualité (selon le paradigme utilisé : possibilité, probabilité, croyance) de cette comparaison.

Plus récemment, une nouvelle série d'indices visant à corriger certains cas de rangements est apparue, fondée cette fois sur des mesures de distances entre quantités floues. Ainsi, (Abbasbandy et Asady, 2006), puis (Abbasbandy *et al.*, 2013) rangent les nombres flous selon leur distance à un nombre flou servant d'origine. Cette distance est finalement rendue orientée pour permettre le classement de nombres négatifs. Cette méthode est une correction de celle de Chu et Tsao (2002) et Wang et Luo (2009) qui s'appuie sur l'aire entre le centre de gravité de la quantité floue et le point d'origine (0, 0). Suivant une approche très similaire, Asady (2011) montrent que le point qui minimise cette distance est le milieu de l'axe moyen de cette quantité, ce qui leur permet de rendre chaque quantité non floue¹⁰ puis de les trier selon l'abscisse de ce point. On peut remarquer que cette mesure est équivalente au quatrième indice de Yager. Enfin, Abbasbandy et Hajjari (2009) introduisent une valeur de *magnitude* synthétisant une indication sur l'aire et la position d'une quantité floue.

Remarques

10. On parle de défuzzification

Bien qu'il existe une grande quantité de méthodes de tri de quantités floues, elles produisent parfois des résultats équivalents. Ainsi, les méthodes de Fortemps et Roubens (1996), Liou et Wang (1992)¹¹, Choobineh et Li (1993)¹² et le second indice de Yager (1981) donnent les mêmes ordonnancements. Sur les exemples fournis par Bortolan et Degani (1985), les mesures de Asady (2011) et Fortemps et Roubens (1996) donnent les mêmes valeurs dès lors que les nombres flous sont positifs. De même, Kołodziejczyk (1986) et Tseng et Klein (1989) fournissent des résultats équivalents (voir pour cela (Chen et Lu, 2001b)).

Propriétés raisonnables d'un tri

L'évaluation de méthodes de tri est classiquement fait à partir d'un ensemble de cas dont l'ordre "naturel" est établi. Cette justification par l'exemple ne prouve cependant pas la cohérence de la méthode évaluée, ce qui en fait sans doute une des causes de l'accumulation de mesures visant à corriger les cas incohérents des précédentes. Pour pallier ce problème, (Wang et Kerre, 2001) ont dressé une liste de sept propriétés d'ordre *raisonnables* traduisant ce que l'on peut considérer comme un ordre naturel. Le principe est simple : plus une mesure vérifie de propriétés, plus elle est jugée satisfaisante.

Les trois premières propriétés portent sur les propriétés d'ordre classiques : réflexivité, antisymétrie, transitivité. La quatrième assure que la méthode se comporte comme un tri classique dans les cas où les quantités sont séparées : si $\text{supp}(A) < \text{supp}(B)$, alors $A \preceq B$. La cinquième stipule que la relation d'ordre entre deux quantités floues est indépendante des autres quantités du problème. La sixième assure la conservation de la relation d'ordre par l'addition. Enfin, la septième propriété conserve la relation d'ordre par la multiplication de quantités floues positives¹³ Les auteurs ont évalué la majorité des mesures antérieures à 2001 et ont remarqué que seule la méthode de Adamo (1980) respectait toutes les propriétés. Le second indice de Yager (1981), la méthode de Kerre (1982) et de Chen (1985) échouent dès la règle quatre. Étant donné leur nature, fondée sur des comparaisons deux à deux, la majorité des méthodes de la troisième catégorie ne respectent pas la propriété 5, à l'exception notable de la proposition de Kołodziejczyk (1986).

Conclusion sur les méthodes de tri

Nous avons dressé ici une liste non exhaustive des méthodes de rangement de quantités floues. Pour chacune, il existe des cas dans lesquels elles créent des rangements contraires à l'intuition humaine. Toutefois, ces cas problématiques ne se rencontrent pas nécessairement dans le cas spécifique du rangement d'instantanés et de durées floues. En outre, on peut constater que les méthodes de la troisième catégorie offrent la possibilité de quantifier la relation de précédence. En effet, la notion de dominance est une grandeur qui peut être assimilée à un degré d'antécédence, comme cela a déjà été effectué par De Runz *et al.* (2010).

Le choix d'une méthode de tri dépend donc du cas d'étude, et des quantités floues

11. Si le paramètre $\gamma = 0.5$.

12. Si les sous-ensembles considérés sont convexes et normaux.

13. Plus précisément, si A, B, AC, BC sont 4 quantités floues avec $C \geq 0$ alors $A \succeq B$ implique que $AC \succeq BC$, où $A \succeq B$ désigne le fait que A a un rang au moins égal à B .

mises en jeu. Nous allons donc commencer par présenter la modélisation des temporalités des objets manipulés, ce qui nous permettra d'effectuer ce choix.

4.4.2 Choix des méthodes de tri temporel et d'antécédence adaptées

Comparer observations et sources par deux méthodes de tri flous complémentaires

Les temporalités des sources géohistoriques comme celles des observations sont des quantités floues normalisées qu'il faut pouvoir ordonner. En effet, la position temporelle des observations et des sources les unes par rapport aux autres est une information essentielle si l'on veut reconstituer des transformations qui suivent la flèche du temps. Avec notre modélisation, nous avons introduit une relation de dépendance entre la source et l'observation. Cette dépendance est également temporelle, puisque les observations ne sont valides que tant que leur source est également valide. Cette dépendance permet de répartir l'objectif de la comparaison temporelle entre observations en deux sous-objectifs complémentaires :

- la création d'un ordre temporel total entre les différents *snapshots* entre lesquels les transformations doivent être identifiées, ce qui revient à ordonner les sources géohistoriques,
- la détermination d'une méthode de mesure du degré d'antécédence permettant de comparer plus finement la temporalité valide de deux observations issues de *snapshots* différents ;

Ainsi, un ordre total entre *snapshots* permet de connaître, pour deux observations, s'il existe d'autres sources géohistoriques entre celles auxquelles elles appartiennent et donc s'il existe d'autres ressources contenant potentiellement des états intermédiaires entre ces deux observations. Le degré d'antécédence permet alors de quantifier l'écart temporel entre deux observations.

Ainsi, la première méthode doit être capable de répondre à des questions telles que, pour deux observations O_1 et O_2 : "Combien existe-il d'autres descriptions de l'espace entre O_1 et O_2 ?". La seconde méthode doit permettre quant à elle de répondre à des requêtes du type "A quel point O_2 a-t-elle été observée après O_1 ?".

Afin de distinguer les deux méthodes, la relation \preceq est utilisée pour indiquer l'ordre entre intervalles temporels.

La notation $ant(A, B)$ désigne quand à elle le degré d'antécédence du temps valide A de O_1 sur le temps valide B de O_2 .

Règles sur tri de temps flous

La majorité des méthodes de tri que nous avons présentées sont adaptées à la comparaison de nombres flous, pour lesquels le problème revient à déterminer une relation de "grandeur". Or, dans le cas du temps, la sémantique associée au tri est quelque peu différente. Il s'agit cette fois d'une relation du type "avant/après", ce qui impose quelques nouvelles propriétés nécessaires pour une méthode de tri de temps flous. Par exemple, l'ordre de deux quantités floues $A = (0, 1, 2, 3)_{LR}$ et $B = (1, 1.5, 2)_{LR}$ est, pour la plupart des méthodes, $B < A$. Or, si l'on parle de temps, A est tout autant avant B qu'il n'est après.

Nous introduisons donc quelques règles supplémentaires sur la méthode de tri et

la méthode de comparaison de deux temps flous afin qu'elles respectent un positionnement temporel intuitif. De plus, nous considérons qu'il est nécessaire que les méthodes choisies vérifient à minima les 5 premières propriétés raisonnables d'un ordre flou qui sont les plus importantes pour assurer un tri cohérent avec l'intuition humaine.

Règle R_1

La première règle est qualitative, et suit un principe de prudence. En effet, il existe de nombreux cas pour lesquels l'intuition est mise en défaut (par exemple, lorsque les deux quantités floues à trier sont presque égales, ou lorsque leur sens réel est difficile à saisir). La définition d'un tri intuitif peut alors varier d'une personne à l'autre. Nous considérons, pour notre cas précis, qu'il vaut mieux en cas de doute conserver une égalité (c'est à dire la coexistence de deux *snapshots*) plutôt que définir un ordre définitif. En effet, il est toujours possible de raffiner l'égalité au niveau des observations grâce au degré d'antécédence, ce qui n'est plus possible si un ordre est définitivement fixé.

Règle R_2

La seconde règle correspond à l'exemple donné précédemment et permet de traiter de temps flous avec une relation "avant" et "après". Pour deux temps flous $A, B \in \tilde{\mathcal{T}}$

$$(\bar{x}_A, \bar{y}_A) = (\bar{x}_B, \bar{y}_B) \implies ant(A, B) = ant(B, A) \wedge A \sim B$$

où (\bar{x}_A, \bar{y}_A) désigne le centre de gravité de A .

Règle R_3

Cette règle est établie pour la mesure du degré d'antécédence. Pour A, B deux temps flou sur T :

$$ant(A, B) = -ant(B, A) \in [0, a_{max}]$$

où $a_{max} > 0$ est la valeur d'antécédence maximale possible entre deux temps flous. Nous considérons donc uniquement des mesures d'antécédences normalisées.

Règle R_4

La dernière règle découle directement de R_2 et définit un "statut-quo" temporel : lorsque deux observations sont faites à la même période, aucune des deux n'est à priori antécédente à l'autre.

$$A \sim B \implies ant(A, B) = ant(B, A) < a_{max}$$

Choix d'une méthode de tri

Nous avons comparé les résultats de 33 méthodes abordées dans la section 4.4 avec un ordre temporel de référence créé manuellement et jugé intuitif, sur 26 exemples de configuration de nombres flous issus de (Chen et Lu, 2001b). L'ensemble des résultats de ces mesures sont fournis dans l'annexe H Parmi ces méthodes, 4 ont donné des résultats cohérents avec le tri manuel : le troisième indice de Yager (1981)

(Yager F3), la mesure *d'intégrale totale* de Kim et Park (1990)¹⁴, la méthode de Fortemps et Roubens (1996) et le tri *minimization distance* de Asady (2011). L'intégrale totale ne respecte cependant ni la cinquième propriété des relations d'ordre flou, ni la règle R_1 . La méthode de Fortemps et Roubens (1996) donne exactement le même tri et les mêmes valeurs que l'incide de Yager F3 (voir (Chen et Lu, 2001b)). Afin de choisir entre ces deux dernières méthodes et celle de Asady (2011), nous avons adopté le principe du Rasoir d'Occam : à résultats équivalents, la plus simple et aisément compréhensible des méthodes est conservé. Ainsi, nous avons choisi d'utiliser comme méthode de tri des temps flous l'indice F3 de Yager.

Cet indice, que l'on note Y_3 , se définit comme suit :

$$Y_3(A) = \int_0^{h(A)} M(A_\alpha) d\alpha \quad (4.9)$$

où $M(A_\alpha)$ est la valeur moyenne des éléments de A_α . Si A est convexe, on a $A_\alpha = [a_\alpha, b_\alpha]$ et donc :

$$Y_3(A) = \frac{1}{2} \int_0^{h(A)} a_\alpha + b_\alpha d\alpha \quad (4.10)$$

L'indice Y_3 consiste à comparer la position des centres de gravités de chaque α -coupes de deux quantités floues. Il revient donc finalement simplement à trier les quantités floues selon le segment de droite passant par tous les centres de gravités de ses α -coupes. Y_3 respecte toutes les propriétés d'ordres à l'exception de la septième. Il respecte également les différentes règles que nous nous sommes imposé.

Choix d'une mesure d'antécédence

Dans le domaine des SIG historique et archéologique, De Runz (2008) a défini une relation d'antécédence fondée sur l'indice de Kerre. Il est également possible d'utiliser toutes les méthodes de tri fondées sur la notion de dominance. Nous avons évalué l'ensemble des méthodes de tri fondées sur la notion de dominance en regard des règles que nous nous sommes fixés. Plus précicément, nous avons évalué les méthodes de De Runz (2008), Kołodziejczyk (1986), Chen et Lu (2001b) et Dubois et Prade (1983). Ces méthodes on été évaluées à partir de 13 situations de dominance entre 2 quantités floues présentées dans (Chen et Lu, 2001b). L'annexe I présente quelques unes des situations critiques de la mesure d'antécédence, pour lesquelles au moins une mesure ne respecte pas les règles que nous nous sommes fixés. Seule la méthode de Kołodziejczyk (1986) respecte l'ensemble de nos règles. La méthode de (De Runz, 2008) échoue lorsque le maximum flou est égal à l'un des deux ensembles. Les autres méthodes ne respectent pas les règles R_2, R_3 et R_4 . Nous choisissons donc finalement la méthode de Kolodziejczyk, dont l'auteur propose deux formes (K_1, K_2). Nous nous appuyons ici sur la mesure K_1 . Pour deux temporalités floues A_i et A_j , nous posons la mesure d'antécédence issue de (Kołodziejczyk, 1986) comme suit :

$$ant(A_i, A_j) = K_1(A_i, A_j) = \frac{d_H(\underline{A}_i, \widetilde{\max}(\underline{A}_i, \underline{A}_j)) + d_H(\overline{A}_i, \widetilde{\max}(\overline{A}_i, \overline{A}_j))}{d_H(\underline{A}_i, \underline{A}_j) + d_H(\overline{A}_i, \overline{A}_j)} \quad (4.11)$$

où d_H est la distance de Hamming entre deux sous-ensembles flous défini par :

$$d_H(A, B) = \int_{\mathbb{R}} |\mu_A(x) - \mu_B(x)| dx \quad (4.12)$$

14. Avec un paramètre $k = 0.5$

Cas d'exemple : temps valides des sources parisiennes

La figure 4.15 présente pour illustration de la méthode de tri et de la mesure d'antécédence leur application aux temps valides des sources géohistoriques sur Paris.

ant(A,B)	Verniquet	Vasserot	Jacoubet	Municipal	
Verniquet	0.5	1	1	1	
Vasserot	0	0.5	0.6	1	Verniquet < Vasserot < Jacoubet < Municipal
Jacoubet	0	0.4	0.5	1	$\Upsilon_3(\mathbf{A})$ 1789,5 1826,7 1831,3 1888
Municipal	0	0	0	0.5	

FIGURE 4.15 – Application des méthodes de comparaison de temps flous aux temporalités des sources parisiennes.

4.5 Conclusion

Dans ce chapitre, nous avons introduit un schéma conceptuel général décrivant les relations entre les traces d'entités géohistoriques perçues au sein de sources géohistoriques et retranscrites sous la forme d'observations au sein de *snapshots* localisés dans le temps.

Nous avons spécialisé ce schéma pour le cas particulier de sources cartographiques vectorisées. Cette implémentation s'appuie sur les standards actuels de représentation de l'information géographique et permet en outre une représentation incertaine des temps valides des observations et des sources cartographiques par le biais de sous-ensembles flous.

Enfin, nous avons réalisé un benchmark des méthodes de comparaison et de tri d'ensembles flous proposés dans la littérature scientifique afin de sélectionner une méthode d'ordonnement et une mesure d'antécédence adaptées au cas de sous-ensembles flous représentant des temporalités incertaines. Ces deux méthodes seront mises à profit lors de la reconstitution des transformations entre observations.

Nous disposons désormais d'un ensemble de *snapshots* vecteurs décrivant l'espace urbain à différentes temporalités. Il nous faut maintenant nous doter d'une structure permettant de décrire des transformations entre les observations peuplant ces *snapshots*.

Chapitre 5

Un graphe pour modéliser les transformations de l'espace

Objectifs :

- Adapter un modèle de graphe spatio-temporel de la littérature au cas de l'information géohistorique afin de modéliser les relations de filiation entre les observations stockées dans la base de données spatiale et temporelle et permettre d'analyser les transformations de l'espace que ces relations traduisent.

Sommaire

5.1	Modèle de graphe spatio-temporel de Del Mondo	211
5.1.1	Notions préalables : entité, identité et filiation	212
5.1.2	Notations préalables	213
5.1.3	Relations entre entités dans le modèle de graphe spatio-temporel	214
5.1.4	Graphe spatio-temporel	216
5.1.5	Instanciation du graphe spatio-temporel dans une base de données relationnelle	218
5.2	Adaptation du graphe spatio-temporel aux observations géo- historiques	221
5.2.1	Redéfinitions et nouvelles notations de base	222
5.2.2	Relations entre observations	223
5.2.3	Spécificités des relations entre observations	224
5.2.4	Graphe géohistorique	228
5.2.5	Contraintes sur le graphe géohistorique	228
5.3	Processus spatio-temporels	230
5.4	Conclusion	236

Nous disposons maintenant d'une base de données spatiale et temporelle permettant de stocker un ensemble d'observations géohistoriques représentées par des *features* vectorielles, structurées par sources géohistoriques et localisées dans le temps de façon imprécise. Cette base de données ne modélise toutefois pas les transformations de ces observations au cours du temps. De plus, pour qu'il soit possible d'analyser ces transformations, nous devons disposer d'un ensemble d'outils permettant d'effectuer des requêtes sur ces transformations. Nous devons donc passer d'un modèle de données spatiales et temporelles à un modèle véritablement spatio-temporel. L'objectif de ce chapitre est donc de proposer une approche permettant de modéliser et de manipuler les transformations entre observations.

Pour atteindre cet objectif, nous nous appuyons sur un modèle de graphe spatio-temporel proposé par Del Mondo (2011). La présentation de ce modèle occupe la section 5.1. Nous proposons par la suite en section 5.2 une adaptation de ce modèle au cas particulier des observations géohistoriques. Cette adaptation, nommée **graphe géohistorique**, permet de modéliser des observations structurées par sources et dont le temps valide est imprécis. Enfin, nous proposons une modélisation formelle d'un ensemble de processus spatio-temporels pertinents pour l'analyse des transformations de l'espace urbain pouvant être inférés depuis un graphe géohistorique.

5.1 Modèle de graphe spatio-temporel de Del Mondo

Dans cette section, nous présentons les éléments du modèle de graphe spatio-temporel de Del Mondo (2011) sur lesquels nous nous appuyons pour construire notre modèle de graphe adapté pour la représentation des transformations entre observations géohistoriques.

Le modèle de graphe spatio-temporel est une structure abstraite permettant de représenter les évolutions d'*entités* localisées dans le temps. Dans ce but, Del Mondo introduit tout d'abord un ensemble d'axiomes permettant d'identifier différents types de relations entre entités nécessaires à la description d'un modèle spatio-temporel. Sont également introduites un ensemble d'opérations ensemblistes permettant de manipuler un graphe spatio-temporel et de fonctions permettant de le parcourir.

Le modèle de graphe spatio-temporel est ensuite instancié au sein d'une base de données relationnelle. La base de données spatio-temporelle qui en résulte, nommée *modèle de données* par son auteur, permet de matérialiser un ensemble d'entités représentées selon le modèle de graphe spatio-temporel. Enfin, deux ensembles de contraintes sont définis sur cette base de données afin d'en assurer la cohérence. Le premier ensemble de contraintes impose certains choix de conception sur les entités stockées. Les autres contraintes sont propres au type de phénomène représenté dans la base de données et permettent de contraindre les types de transformations entre entités que la base peut représenter.

Nous présentons ici les aspects principaux du modèle de graphe spatio-temporel général ainsi que la modélisation de celui-ci et des contraintes qui l'accompagnent au sein d'une base de données relationnelle. Certains aspects du modèle, en particulier les fonctions de jointure entre graphes, ne sont pas abordées. Nous réutilisons les notations d'origines autant que faire se peut.

5.1.1 Notions préalables : entité, identité et filiation

Entité

L'entité est l'élément de base du modèle de graphe spatio-temporel. Une entité est un objet générique, possédant une étendue spatiale et localisé temporellement. Il faut toutefois distinguer l'entité du modèle d'un objet géographique ou d'une entité du monde réel. Il s'agit en effet d'une entité au sens des bases de données, c'est-à-dire un objet stocké dans une base de données et représenté par un ensemble d'attributs. Ici, l'entité désigne une représentation d'un état d'une entité du monde réel enregistré à un instant précis. Si l'entité du monde réel persiste pendant une certaine durée, plusieurs entités du modèle spatio-temporel peuvent la représenter¹. Toute entité du graphe spatio-temporel possède une identité. Une entité est finalement dans le modèle spatio-temporel un objet formé d'un ensemble d'attributs et d'une identité, correspondant à l'état d'une entité du monde réel à un instant donné. L'entité du modèle de Del Mondo est donc celle des bases de données, à différencier de l'entité du monde réel au sens géographique du terme. Dans ce chapitre, le terme d'*entité* seul désignera toujours les objets manipulés par le graphe. Nous nommerons explicitement *entité du monde réel* les objets du monde réel dont une ou plusieurs représentations sont stockées dans le graphe.

Identité

L'identité est un concept fondamental de toute approche traitant d'entités (au sens large, voir (Mathian, 2014, pp.20–21)) localisées dans le temps, évoluant et se transformant. L'identité d'une entité est **une caractéristique qui distingue l'entité de toutes les autres** (Hornsby et Egenhofer, 2000). L'identité est en général difficile à définir et est très dépendante de la nature de l'entité concernée. Dans le modèle de graphe spatio-temporel de Del Mondo, l'identité des entités est considérée comme connue a priori.

Filiation

Des entités sont affiliées lorsque leurs identités sont liées. Les entités du graphe pouvant représenter plusieurs états successifs d'une même entité du monde réel, il peut exister plusieurs entités du graphe avec la même identité mais représentant une entité du monde réel existant à des instants différents. Dans ce cas, il y a égalité entre les identités de ces différentes entités. Une autre catégorie de filiation concerne la généalogie, et exprime le fait qu'une entité est le "parent" ou "l'enfant" d'une autre. Cette relation dépend également de l'identité, et donc d'un certain point de vue. Par exemple, on peut considérer que l'actuel musée d'Orsay "descend" de l'ancienne gare de l'ouest si l'identité porte sur la structure du bâti. Il n'y a en revanche pas de filiation si l'identité concerne la fonction du bâti (gare et musée ne sont à priori pas des concepts sémantiquement proches).

1. Dans la taxonomie des dynamiques de l'espace de (Grenon et Smith, 2004), il peut s'agir tout autant d'entités *SPAN*, que *SNAP*, aussi appelées perdurantes et endurantes (Mathian, 2014, p.32). Il peut en effet s'agir de la représentation d'une entité n'ayant pas de composante temporelle et donc restant "statique" dans le temps, ou bien d'une entité perdurante représentée pendant une période de stabilité.

5.1.2 Notations préalables

Le modèle de graphe spatio-temporel s'appuie sur la notion de *Dynamic Set* introduite par Stell (2003). Un *Dynamic Set* $X : T \rightarrow Rel$ décrit un ensemble de **relations** Rel pouvant exister au sein d'un **ensemble d'entités** σ localisées temporellement grâce à un **domaine temporel** (T, \leq) . Un *Dynamic Set* décrit donc les relations entre des entités évoluant dans le temps. Il s'appuie en particulier sur une **relation support** générique représentant les relations pouvant exister entre entités localisées dans le temps. Nous précisons dans les paragraphes ci-dessous le domaine temporel, la relation support et deux fonctions définies sur un *Dynamic Set* permettant de manipuler des entités localisées temporellement. Enfin, nous introduisons une fonction d'**identification d'une entité**, décrivant l'identité d'une entité, qui ne provient pas du modèle de Del Mondo mais nous permet d'écrire plus simplement certaines relations particulières dans la suite de ce document.

Domaine Temporel (T, \leq)

Le domaine temporel (T, \leq) est défini de manière similaire à celui présenté dans le chapitre 4 (voir 4.3.2). Le modèle s'appuie sur une définition linéaire du temps, discrétisé selon une granularité quelconque. L'ensemble fini des instants $T = \{t_0, \dots, t_n\}$, où $n \in \mathbb{N}$, est associé à un ordre partiel \leq recréant la flèche du temps. Le modèle n'impose aucune restriction sur la modélisation des instants (points, dates, etc.). Deux propriétés particulières sont définies sur le domaine temporel :

- T étant fini, il y a donc un instant initial t_0 et un instant final t_n . Leur valeur dépend de l'application.
- \leq est un ordre discret : deux instants t_i et t_j sont consécutifs si $j = i + 1$.

Relation support $X(t_i, t_j)$

Dans le *Dynamic Set*, la relation support notée $X(t_i, t_j)$ définie pour tout couple d'instant t_i, t_j tels que $t_i \leq t_j$ décrit toute relation pouvant être définie entre deux entités localisées temporellement. Il s'agit donc d'une relation générique à partir de laquelle on peut dériver des types de relations propres à un cadre applicatif.

Localisation temporelle $f(a) = t$

La fonction $f : \sigma \rightarrow T$ décrit la localisation temporelle de chacune des entités du modèle. Elle permet de connaître l'instant d'existence **unique** de chacune de ces entités. L'entité du modèle de graphe spatio-temporel correspond donc à un instantané d'une entité du monde réel.

Snapshot spatio-temporel $X(t) = \{a | f(a) = t\}$

La fonction $X : T \rightarrow \sigma^k$ renvoie l'ensemble des k entités existant à un instant t donné. Cette fonction construit donc un snapshot au sens des modèles spatio-temporels éponymes, c'est-à-dire un ensemble d'entités décrivant un état de l'espace figé à un instant précis.

Identification d'une entité $i(a)$

Toute entité $a \in \sigma$ possède une identité. Dans le modèle de graphe spatio-temporel, cette identité désigne un **ensemble de caractéristiques** qui distingue l'entité a de toutes les autres entités **existant au même instant**.

Plusieurs entités peuvent donc posséder la même identité si elles sont localisées à différents instants. Cela traduit le fait que l'entité du modèle n'est en réalité pas une représentation d'une entité du monde réelle évoluant dans le temps, mais un état de cette entité à un instant unique. Plusieurs états successifs de la même entité du monde réel aboutissent donc à l'existence de plusieurs entités ayant la même identité dans le graphe spatio-temporel.

Nous introduisons ici une notation générale de l'identité qui nous est utile pour le reste de la présentation du modèle. Sans le définir plus précisément, nous notons \mathcal{I} l'ensemble des identités d'un ensemble d'entités σ .

Nous posons alors la **fonction d'identification d'une entité** $i : \sigma \rightarrow \mathcal{I}$ qui renvoie l'ensemble des caractéristiques définissant l'identité d'une entité.

5.1.3 Relations entre entités dans le modèle de graphe spatio-temporel

Dans le modèle de graphe spatio-temporel, la relation support générique $X(t_i, t_j)$ est spécialisée en un ensemble de relations ρ pour représenter des **relations spatiales, spatio-temporelles** ou décrivant un lien de **filiation** entre entités. Ainsi, $\rho = \{\rho_s, \rho_{st}, \rho_f\}$ où ρ_s désigne une relation spatiale, ρ_{st} une relation spatio-temporelle et ρ_f une relation de filiation.

On note alors $a \rho b$ l'existence d'une relation entre deux entités a et b .

Les sémantiques des trois types de relation définies pour tout couple d'entités (a, b) tel que $f(a) \leq f(b)$ sont décrites ci-dessous. Chaque relation est accompagnée par une représentation graphique issue du modèle de graphe spatio-temporel et résume les contraintes sur l'existence de cette relation entre les deux entités a et b .

Relation spatiale $a \rho_s b$

Il y a relation spatiale entre a et b lorsque les deux entités existent au même instant et qu'il y a un **connexion spatiale** entre elles. Pour exprimer les différentes configurations de connexion spatiale possibles, Del Mondo s'appuie sur l'algèbre RCC8 (*Region Connexion Calculus*) Randell *et al.* (1992) qui distingue sept cas de connexions possibles et un cas de déconnexion. Par exemple, la France et l'Allemagne sont deux pays connectés spatialement² car ils partagent une frontière commune. La connexion spatiale entre a et b étant également une relation, elle est notée $a s b$.

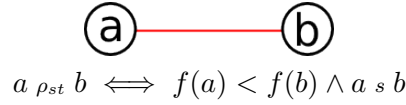
$$\begin{array}{c}
 \textcircled{a} \text{---} \textcircled{b} \\
 a \rho_s b \iff f(a) = f(b) \wedge a s b
 \end{array}$$

Relation spatio-temporelle $a \rho_{st} b$

Cette relation exprime l'évolution des relations spatiales entre deux entités. L'existence de l'entité a précède donc celle de l'entité b . La relation spatiale est

2. Il s'agit de la relation EC *externally connected* du RCC8.

restreinte aux connexions spatiales. Or, une relation de connexion est souvent insuffisante pour décrire un processus spatio-temporel (le croisement de deux avions n'est par exemple pas régi par une relation de connexion spatiale). Comme l'auteur l'indique, rien n'interdit dans le modèle de considérer des types relations spatiales supplémentaires.

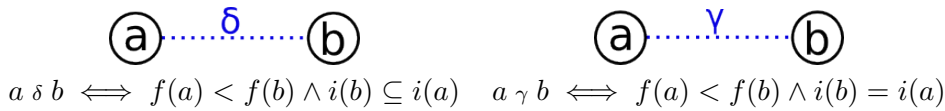
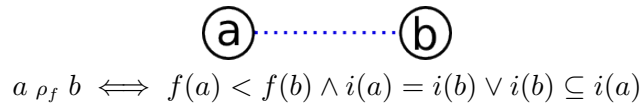


Relation de filiation $a \rho_f b$ et spécialisations (δ, γ) liées à l'identité

Il y a filiation si l'existence de l'entité b dépend de l'existence de l'entité a , et que les deux entités ne coexistent pas. Dans le modèle, cela se traduit par le fait que l'identité de b est liée à celle de a .

Deux cas peuvent alors se présenter :

- si une partie de l'identité de b est héritée de celle de a , cela signifie qu'il existe une relation généalogique entre les deux entités (du type parent/enfant). Par exemple, la division d'une parcelle en deux peut être vue de cette façon : les parcelles issues du découpage héritent d'une partie de l'identité de la parcelle d'origine (l'emprise spatiale notamment), mais ne sont pas pour autant des parcelles identiques à celle d'origine. Cette relation est nommée **dérivation** et est notée $a \delta b$,
- si les identités de a et b sont identiques, cela signifie qu'il s'agit d'une même entité du monde réel. Celle-ci perdure pendant plusieurs instants, mais du fait de la modélisation du temps en instants elle se trouve représentée plusieurs fois. a et b entretiennent alors une relation de **continuation** notée $a \gamma b$.



Finalement, la figure 5.1 résume l'ensemble des relations considérées dans le modèle de graphe spatio-temporel.

Contraintes sur la transmission de l'identité

La filiation de type continuation peut être vue comme un lien de transmission de l'identité. Or, dans le modèle, cette transmission fait l'objet de deux restrictions (voir la figure 5.2) :

- une entité a doit être en relation de continuation avec au plus une autre entité. En effet, une telle possibilité poserait deux problèmes. D'une part, si a est en relation de continuation avec b et c avec $f(b) = f(c)$, il y aurait alors $i(b) = i(c)$, ce qui est contraire à la définition de l'identité. D'autre part,

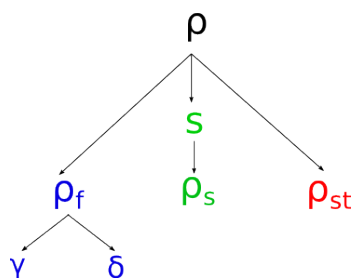


FIGURE 5.1 – Hiérarchie des relations considérées dans le modèle de graphe spatio-temporel.

ceci signifierait que la totalité de l'identité de a est transférée à deux autres entités, situation non autorisée par le modèle.

- s'il existe $a \gamma b$, alors b ne peut être issue d'aucune autre entité. Le modèle considère en effet qu'un tel cas est incompatible avec l'idée de continuation.



FIGURE 5.2 – Situations interdites dans le graphe spatio-temporel. Figure d'après Del Mondo (2011).

5.1.4 Graphe spatio-temporel

Définition

Les entités et les relations spatiales, spatio-temporelles et de filiation qui existent entre ces entités forment un graphe spatio-temporel $G_{ST} = (\sigma, E_{ST}, T)$, avec $E_{ST} = \{(a, b) | a \rho b\}$ et $a, b \in \sigma$. Ce graphe global peut être décomposé en trois sous-graphes :

1. Un **sous-graphe spatial** $G_s = (\sigma, E_s, T)$ tel que $E_s = \{(a, b) | a \rho_s b\}$.
2. Un **sous-graphe spatio-temporel** $G_{st} = (\sigma, E_{st}, T)$ tel que $E_{st} = \{(a, b) | a \rho_{st} b\}$.
3. Un **sous-graphe de filiation** $G_f = (\sigma, E_f, T)$ tel que $E_f = \{(a, b) | a \rho_f b\}$.

Le graphe G_{ST} s'écrit donc $(\sigma, E_s \cup E_{st} \cup E_f, T)$.

Un exemple de graphe spatio-temporel est présenté en figure 5.3. Il représente les transformations d'un ensemble de quelques parcelles cadastrales à Paris entre trois instants. Toutes les relations spatio-temporelles notées en rouge ne sont pas représentées pour des questions de lisibilité. En effet, toutes les relations de filiation devraient en réalité s'accompagner de relations spatio-temporelles, les parcelles successives étant superposés. L'identité des différentes parcelles cadastrales est connue a priori, ce qui permet d'inférer les relations de type "continuation" et "dérivation".

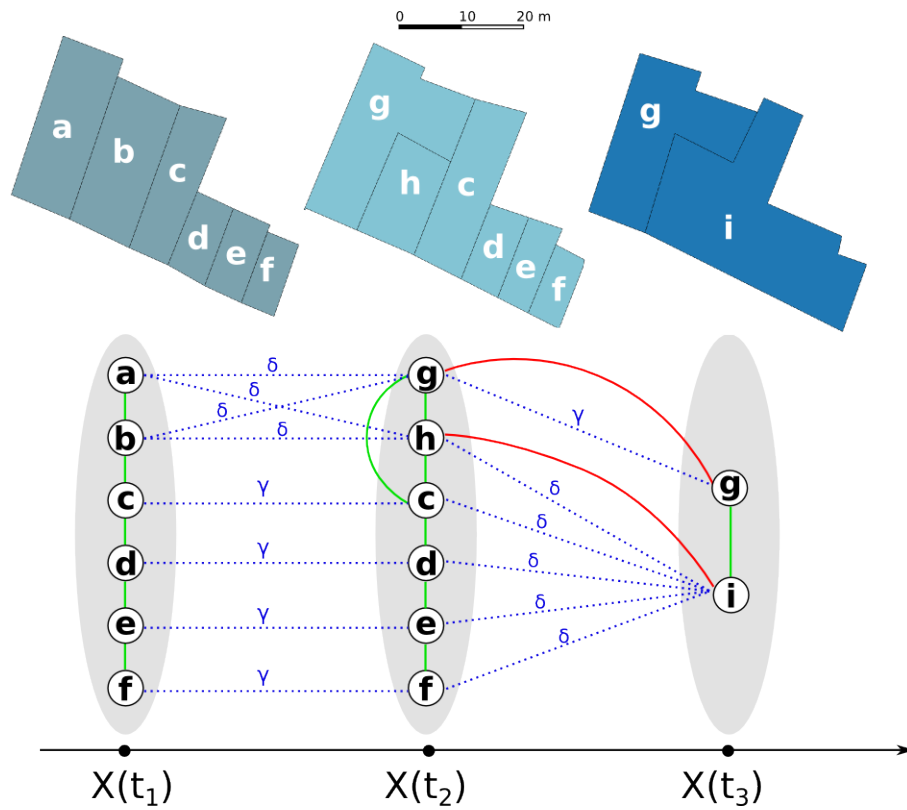


FIGURE 5.3 – Un exemple de graphe spatio-temporel représentant les transformations de parcelles cadastrales à trois instants. Toutes les relations spatio-temporelles ne sont pas représentées.

Fonctions de voisinage

Le modèle définit plusieurs fonctions de voisinage dans les sous-graphes formant G_{ST} qui permettent de parcourir celui-ci. Le sens de ces fonctions dépend du sous-graphe de G_{ST} sur lequel elles s'appliquent. Par exemple, une fonction de voisinage dans le sous-graphe de filiations G_f décrit des relations de descendance et de succession. Cela revient donc à considérer les arêtes du graphe spatio-temporelle comme orientées.

Tout comme les différentes relations du graphe qui sont des spécialisations de la relation générique ρ , les fonctions de voisinage ont une forme générale $\rho(a)^{+/-}$, $a \in \sigma$. Les fonctions $\rho(a)^+$ et $\rho(a)^-$ renvoient respectivement les entités **succédant** et **précédant** a^3 .

Elles se trouvent définies pour tout couple d'entités $a, b \in \sigma$ par :

$$\begin{aligned} \rho(a)^+ &= \{b \in \sigma \mid a \rho b\} \\ \rho(a)^- &= \{b \in \sigma \mid b \rho a\} \end{aligned} \quad (5.1)$$

Ces deux fonctions peuvent être appliquées de manière récursive pour extraire un

3. $\rho(a)^+$ désigne donc les nœuds de G_{ST} connectés à des arêtes sortantes de a et $\rho(a)^-$ à des arêtes entrantes.

voisinage à une distance $d \in \mathbb{N}$ donnée de a dans le graphe.

$$\begin{aligned}\rho(a)^{+d} &= \{b \in \sigma \mid a \rho^d b\} \\ \rho(a)^{-d} &= \{b \in \sigma \mid b \rho^d a\}\end{aligned}\tag{5.2}$$

Où ρ^d représente l'existence d'un chemin de longueur d dans le graphe composé exclusivement d'arêtes d'un type parmi $\{\rho_s, \rho_{st}, \rho_f\}$.

Étant donné que $\rho(a)^+$ est un cas particulier de $\rho(a)^{+d}$ pour $d = 1$, on ne précisera dans les prochaines définitions que le cas le plus général.

Voisinage spatial

Dans le domaine spatial, les notions de succession et de précédence perdent de leur sens. Cependant, les relations spatiales peuvent tout à fait être orientées ; c'est notamment le cas de l'inclusion. Les fonctions de voisinage s'écrivent dans ce cas :

$$\begin{aligned}\rho_s(a)^{+d} &= \{b \in \sigma \mid a \rho_s^d b\} \\ \rho_s(a)^{-d} &= \{b \in \sigma \mid b \rho_s^d a\}\end{aligned}\tag{5.3}$$

Voisinage spatio-temporel

Lorsque le temps intervient, les entités succédant à une autre correspondent à ses **enfants**, tandis que celles la précédant sont ses **parents**. On écrit alors :

$$\begin{aligned}\rho_{st}(a)^{+d} &= \{b \in \sigma \mid a \rho_{st}^d b\} \\ \rho_{st}(a)^{-d} &= \{b \in \sigma \mid b \rho_{st}^d a\}\end{aligned}\tag{5.4}$$

Voisinage en filiations

De même que pour les relations spatio-temporelles, les relations de filiations permettent de naviguer dans l'histoire d'une entité. Cette fois, les entités lui succédant par des relations de filiation sont désignées comme ses **descendants**, et celles la précédant ses **ancêtres**.

$$\begin{aligned}\rho_f(a)^{+d} &= \{b \in \sigma \mid a \rho_f^d b\} \\ \rho_f(a)^{-d} &= \{b \in \sigma \mid b \rho_f^d a\}\end{aligned}\tag{5.5}$$

5.1.5 Instanciation du graphe spatio-temporel dans une base de données relationnelle

Le graphe spatio-temporel est une structure abstraite permettant d'inférer un certain nombre de connaissances sur l'évolution d'entités localisées dans l'espace et le temps grâce à un ensemble de fonctions prédéfinies. Pour que ces opérations soient possible, il est nécessaire de stocker les entités et d'instancier les différentes fonctions et prédicats constituant le modèle général dans une base de données. Dans la proposition de Del Mondo, le graphe spatio-temporel est instancié sous la forme d'un schéma de données relationnel décrivant l'organisation des entités stockées et nommé "modèle de données". Pour effectuer cette instanciation, certains éléments du modèle doivent être adaptés à un schéma relationnel. En particulier :

- il faut disposer d'une définition de l'identité des entités stockées,

- les entités doivent respecter un certain nombre de contraintes d'intégrité afin qu'il soit possible d'extraire des connaissances sur leurs évolutions en utilisant les fonctions et prédicats définis dans le modèle de graphe spatio-temporel.

Les choix effectués pour répondre à ces deux contraintes sont rapidement présentés ici car ils conditionnent en partie l'adaptabilité du modèle de graphe spatio-temporel aux observations géohistoriques.

Notons que dans les exemples donnés dans le reste de la présentation du modèle de Del Mondo, on ne considère dans un souci de simplicité d'écriture que des entités situées à des instants consécutifs. Ceci n'est cependant pas imposé par le modèle.

Représentation de l'identité

Pour inférer les relations présentées dans la section précédente, il faut disposer d'une représentation de l'identité des différentes entités stockées. Dans le modèle de données, les entités sont représentées par un prédicat $O(id, t, u_1, \dots, u_n, g)$ où, sans entrer dans les détails :

- id est un attribut représentant l'identité,
- t est l'instant d'existence de l'entité,
- $\{u_1, \dots, u_n\}$ est un ensemble d'attributs quelconque dépendant de la nature de l'entité,
- g est la composante spatiale de l'entité, représentée par sa géométrie.

Dans le modèle de données, l'identité est donc un attribut connu a priori pour chaque entité, ce qui permet directement d'extraire les relations de continuation. Par exemple, deux entités $O_1(i_1, t_1, \dots)$ et $O_2(i_1, t_2, \dots)$ se trouvent en relation de continuation, leurs identités étant égales et localisées à deux instants successifs.

Il n'est cependant pas possible d'extraire les relations de dérivation. Dans leur cas, il est nécessaire de les représenter explicitement dans la base de données. Une relation de dérivation est exprimée par un prédicat $R(id_1, t_1, \delta, id_2, t_2)$ La figure 5.4 présente un exemple de l'instanciation d'une partie du graphe illustré (pour les parcelles a,b,c,g et h) en figure 5.3.

Contraintes d'intégrité sur la base de données spatio-temporelle

Pour qu'il soit possible d'inférer des connaissances sur les évolutions des entités stockées dans le modèle de données, il faut maintenir une base de données cohérente. Cette cohérence est assurée par un ensemble de contraintes d'intégrité posées sur la base de données. Ces contraintes sont de deux types :

- des contraintes d'intégrité propres au modèle,
- des contraintes propres à un cas d'application, c'est-à-dire des contraintes métier.

Les premières sont les traductions des contraintes posées dans la partie générique, en particulier sur les relations de dérivation et de continuation. Les contraintes concernant la transmission de l'identité (voir le paragraphe 5.1.3) sont transférées dans le modèle de données.

Une contrainte particulière est posée sur l'identité, en plus de la succession temporelle définie par la relation de filiation : **si deux entités ont la même identité, leurs attributs et leur géométries sont identiques.**

Enfin, des contraintes propres au type de phénomène représenté par le modèle de

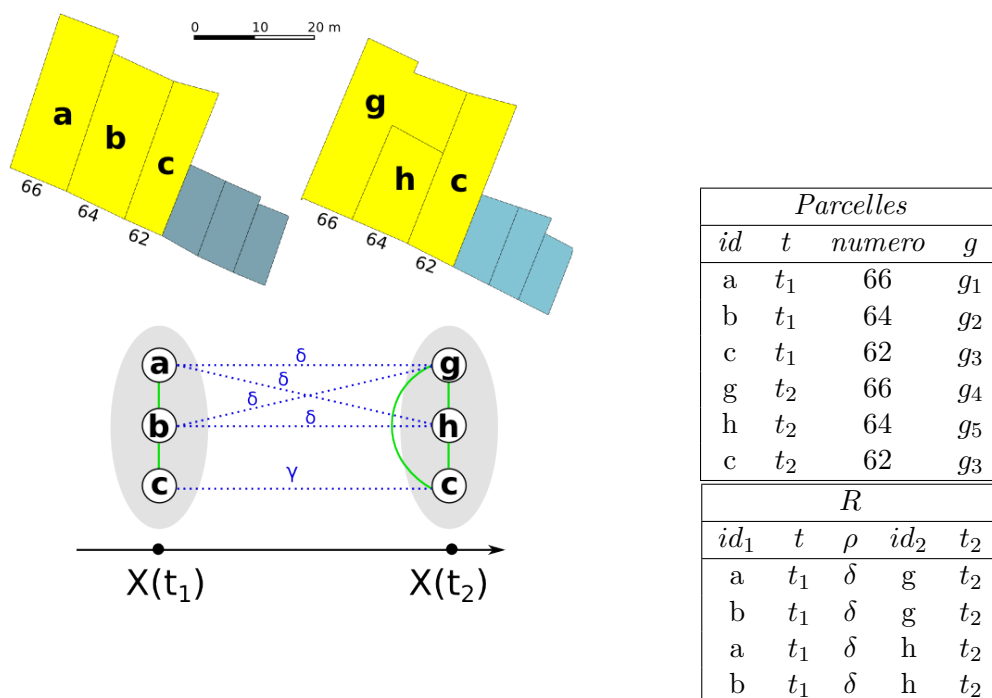


FIGURE 5.4 – Exemple d’instanciation du modèle de graphe spatio-temporel (à gauche) en un schéma de base de données relationnel (à droite).

données peuvent être ajoutées. Ces contraintes, nommées **contraintes sémantiques** correspondent à des processus spatio-temporels et permettent de contrôler les types d’évolutions entre entités que le modèle peut représenter. Del Mondo propose ainsi une modélisation sous forme de contrainte des processus de division, séparation, fusion et annexion que l’on retrouve dans Claramunt et Thériault (1995) et Galton (2004). Par exemple, la contrainte sémantique relative à la fusion de deux entités s’écrit (cf. figure 5.5) :

$$\begin{aligned} & \forall (O_1(id_1, t_1, \dots, g_1) \wedge O_2(id_2, t_2, \dots, g_2) \wedge O_3(id_3, t_3, \dots, g_3) \\ & \wedge R(id_1, t_1, \delta, id_3, t_3) \wedge R(id_2, t_2, \delta, id_3, t_3) \rightarrow g_3 = g_1 \cup g_2 \end{aligned} \quad (5.6)$$

où $g_1 \cup g_2$ désigne l’union géométrique et $g = g'$ l’égalité entre deux géométries (forme, localisation).

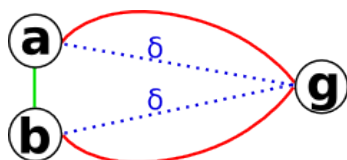


FIGURE 5.5 – Contrainte sémantique décrivant une fusion.

Conclusion sur le modèle de graphe spatio-temporel et son instanciation

Les divers prédicats du modèle que nous avons présentés permettent donc de construire un graphe spatio-temporel à partir d’une base de données relationnelle

respectant un schéma et un certain nombre de contraintes d'intégrité. Les relations définies sur les routes du graphe offrent des outils intéressants pour analyser le graphe que l'on peut extraire d'une telle base de données. Un tel modèle est particulièrement adapté pour la représentation des transformations de l'espace et reste compatible avec l'organisation en *snapshots* spatio-temporels que nous avons proposée dans le chapitre 4. Notre objectif est donc de construire une base de données à partir de laquelle un graphe spatio-temporel puisse être extrait.

Dans notre cas, les entités sont des observations géohistoriques instanciées sous la forme de features et organisées par source géohistorique. Les observations présentent des particularités que les entités du graphe n'ont pas, et pour cette raison il est nécessaire d'adapter le modèle de graphe spatio-temporel au cas des observations géohistoriques. Cette adaptation est l'objet de la section suivante.

5.2 Adaptation du graphe spatio-temporel aux observations géohistoriques

Les observations géohistoriques sont compatibles avec la notion d'entité utilisée dans le modèle de graphe spatio-temporel puisqu'il s'agit de traces d'objets géographiques localisées dans le temps. Trois différences fondamentales existent cependant :

- les observations géohistoriques sont localisées de manière imprécise dans le temps et sont valides pendant des périodes et non des instants,
- les observations sont organisées au sein de sources séparées : il n'y a pas de représentation unique de l'espace qui regrouperait toutes les observations,
- l'identité des observations n'est pas connue par avance (sans prétraitement manuel important tout du moins).

Ces différences ont chacune des conséquences sur l'intégration des observations dans le graphe de Del Mondo. En particulier :

- les temps valides étant des périodes floues, l'utilisation du *Dynamic Set* dans le modèle de graphe n'est pas compatible avec les observations géohistoriques,
- sans identité, l'identification des filiations est impossible. On ne peut donc pas extraire de graphe spatio-temporel facilement à partir des observations.

Pour cette raison, nous présentons dans cette section les adaptations nécessaires pour qu'il soit possible d'extraire un graphe spatio-temporel à partir d'un ensemble d'observations stockées dans une base de données spatiale et temporelle. L'identification des filiations entre les observations de cette base de données est un verrou majeur qui sera, quant à lui, abordé dans le chapitre 6. Dans ce chapitre, nous restons à un niveau abstrait où l'identité peut être utilisée sans être définie précisément. Afin de distinguer le modèle proposé par Del Mondo (2011) de nos adaptations, nous regroupons celles-ci sous le nom de **graphe géohistorique**.

Enfin, notre objectif est de disposer d'un graphe géohistorique à partir duquel peuvent être extraits un certain nombre de processus de transformation propres à l'espace urbain. Nous introduisons donc ici également ces processus et les règles associées qui permettent de les détecter dans un graphe géohistorique existant.

5.2.1 Redéfinitions et nouvelles notations de base

Pour conserver le lien avec le modèle d'origine et faciliter la lecture, nous conservons autant que possible les mêmes notations. Dès lors, toutes les notations communes renverront à leur interprétation pour le graphe géohistorique⁴. Nous conservons ainsi l'ensemble σ qui désigne cette fois l'ensemble des observations géohistoriques.

Dynamic Set pour le cas géohistorique

Le *Dynamic Set* est maintenant $X : \mathcal{T} \rightarrow Rel$ et décrit un ensemble de relations entre des observations géohistoriques valides pendant des périodes mal connues décrites dans un **domaine temporel flou** \mathcal{T} . Ces observations géohistoriques sont organisées selon les **sources géohistoriques** desquelles elles proviennent, ces sources possédant aussi un temps valide. La relation support possède toujours la même sémantique mais porte maintenant sur des temporalités floues. La fonction de **localisation temporelle** s'applique elle aussi sur des temps flous et doit être déclinée pour les sources géohistoriques.

Domaine temporel flou \mathcal{T}

Dans notre cas, la primitive temporelle n'est plus l'instant mais l'intervalle. En conséquence, le domaine temporel devient (\mathcal{T}, \leq) . \mathcal{T} est l'ensemble des temporalités valides des observations constituant σ . De la même façon que t désignait un instant dans le domaine temporel, on note \tilde{t} une temporalité valide⁵. De plus, comme pour T , le domaine temporel flou doit être défini à une certaine granularité. Enfin, \leq est un ordre partiel.

Ensemble des sources géohistoriques S

En plus des observations, nous traitons de sources géohistoriques. Nous ajoutons en plus de σ l'ensemble S des sources géohistoriques.

Source d'une observation géohistorique s

Dans le cas des observations géohistoriques, le sens d'un snapshot diffère de celui du modèle de graphe spatio-temporel. Dans ce dernier, il s'agit de l'ensemble des entités existant à un instant donné (voir le paragraphe 5.1.2). Dans notre cas, un *Snapshot* contient des observations qui se trouvent représentées dans une source. Le *Snapshot* n'est qu'un conteneur d'observations, la dépendance véritablement importante se trouvant entre l'observation et sa source géohistorique. Nous introduisons une nouvelle fonction $s : \sigma \rightarrow S$ qui renvoie la source géohistorique de laquelle provient une observation.

4. Sauf cas particuliers systématiquement précisés.

5. \tilde{t} est un sous-ensemble flou de T , associé à une fonction d'appartenance $\mu_{\tilde{t}}$. Dans un souci de lisibilité, lorsque l'on donnera un temps flou, seul le sous-ensemble sera précisé. Leurs fonctions d'appartenance peuvent être déduites directement du sous-ensemble puisqu'il ne s'agit que de nombres et d'intervalles convexes dont les fonctions L et R sont linéaires.

Localisations temporelles F et f

La fonction de localisation temporelle doit être également redéfinie. En effet, observations et sources ne sont plus reliées à un instant unique. On introduit $F : S \rightarrow \mathcal{T}$ qui renvoie le temps valide d'une source géohistorique, et $f : \sigma \rightarrow \mathcal{T}$ devient une fonction renvoyant le temps valide d'une observation géohistorique.

Identification d'une observation i


La fonction d'identification d'une entité a été définie de façon suffisamment générale pour qu'elle s'applique également aux observations géohistoriques. Cependant, la sémantique de l'identité change : **l'identité d'une observation géohistorique $a \in \sigma$ est un ensemble de caractéristiques qui la distinguent de toutes les autres observations présentes au sein de la même source géohistorique.** Contrairement aux entités du graphe spatio-temporel, plusieurs observations de même identité peuvent être valides au même instant ou à la même période dès lors qu'elles proviennent de sources différentes. En revanche, deux observations avec la même identité ne peuvent cohabiter au sein d'une même source, même si ces observations ont des temps valides différents⁶

5.2.2 Relations entre observations

Certaines composantes de *Dynamic Set* ayant été modifiées, notamment l'identité et le domaine temporel, les relations $\rho = \{\rho_s, \rho_{st}, \rho_f\}$ doivent être redéfinies.

Relation spatiale $a \rho_s b$


Nous conservons ici la même relation de connexion spatiale s que pour le graphe spatio-temporel, c'est-à-dire limitée aux relations de connexion spatiale. Une relation spatiale entre deux observations a et b prises dans σ peut exister si elles sont en connexion spatiale, si elles peuvent avoir été effectuées au même moment et si elles proviennent de la même source géohistorique.

$$a \rho_s b \iff \underbrace{f(a) \cap f(b) \neq \emptyset}_{\text{observation conjointe}} \wedge \underbrace{s(a) = s(b)}_{\text{source commune}} \wedge \underbrace{a s b}_{\text{connexion spatiale}}$$


Relation spatio-temporelle $a \rho_{st} b$

La relation spatio-temporelle garde essentiellement la même forme, mais il ne peut y avoir de relation spatio-temporelle au sein d'une même source. De plus, il peut y avoir une telle relation dès que a est antécédent à b ($ant(a, b) > 0$, cf. chapitre 4) Pour deux observations $a, b \in \sigma$:


6. Cette décision peut être discutable : on peut imaginer en effet une source décrivant plusieurs états d'une même entité du monde réel. Cependant, nous avons considéré jusqu'ici que les sources -les plans plus particulièrement- sont des vues de l'espace urbain et contiennent donc un seul état de cet espace. Certains plans présentant par exemple les projets d'alignement superposés aux rues doivent être traités comme deux versions du plan.

$$a \rho_{st} b \iff \underbrace{a \text{ s } b}_{\text{connexion spatiale}} \wedge \underbrace{ant(a, b) > 0}_{a \text{ peut être observée avant } b} \wedge \underbrace{s(a) \neq s(b)}_{\text{sources différentes}}$$


Relation de filiation $a \rho_f b$

La filiation entre observations possède le même sens que dans le cas d'entités du graphe spatio-temporel. Deux observations a et b sont en relation de filiation si a transmet une partie de son identité à b ou si leurs identités sont égales. Cette relation ne peut toutefois exister qu'entre sources géohistoriques différentes et si b peut être observée après a . Comme nous restons dans le cas général, les opérateurs $=$ et \subseteq sur les identités sont considérés valides.

On a alors : Les relations de dérivation et de continuation peuvent être dérivées

$$a \rho_f b \iff \underbrace{ant(a, b) > 0}_{a \text{ peut être observée avant } b} \wedge \underbrace{s(a) \neq s(b)}_{\text{sources différentes}} \wedge \underbrace{i(a) = i(b) \vee i(a) \subseteq i(b)}_{\text{identités liées}}$$


directement de la relation de filiation : on a alors $(i(a) = i(b))$ pour une continuation et $i(a) \subseteq i(b)$ pour une dérivation.

5.2.3 Spécificités des relations entre observations

Dans cette section, nous présentons les spécificités des observations géohistoriques qui nécessitent de relâcher certaines contraintes du graphe spatio-temporel ou d'introduire de nouveaux mécanismes. Nous redéfinissons également quelques concepts existants avec le graphe spatio-temporel mais ayant une sémantique particulière dans notre cas.

Sauts temporels

Les sources géohistoriques sont une vue filtrée, abstraite et déformée d'un état de l'espace urbain. Dans le cas des sources cartographiques, tous les constituants de cet espace ne sont pas représentés en raison de l'objectif de la carte, des choix de légende, de son échelle et des choix de généralisation cartographique. Pour ces différentes raisons, il peut arriver qu'une source cartographique ne présente aucune trace d'une entité du monde réel qui se trouve pourtant renseignée par une source antérieure et une autre postérieure. Il peut donc exister des "trous" dans la suite de traces d'une même entité du monde réel. Sur les rues de Paris, c'est le cas pour de nombreuses impasses ou petites rues (passages, etc.), qui ne figurent pas sur certains plans en raison de leur trop faible importance par rapport au niveau de détail de la carte.

Le modèle de graphe spatio-temporel offre la possibilité d'extraire des relations de filiation et des relations spatio-temporelles entre des entités existant à des instants successifs mais non consécutifs. Dans notre cas, ceci se transpose dans la possibilité d'identifier de telles relations entre des sources non nécessairement consécutives dans le temps.

Sens rétro-temporel local

Le graphe spatio-temporel est un graphe orienté suivant la flèche du temps. Or, l'utilisation de temps valides flous et par périodes pour les observations géohistoriques crée des chevauchements entre temps valides. Ces recouvrements correspondent à des périodes pendant lesquelles plusieurs observations peuvent être valides sans qu'il soit possible de les ordonner simplement dans le temps. Ceci est directement lié à la nature même des observations qui saisissent pendant une période mal connue l'état d'une entité du monde réel⁷. Finalement, un chevauchement temporel entre deux observations a et b affiliées traduit le fait que pendant une certaine durée il est impossible de savoir si l'entité du monde réel correspondant à a a été observée avant celle correspondant à b ou l'inverse. Dans notre modèle, ceci signifie que l'antécedence temporelle n'est plus stricte avec $ant(a, b) > 0$ et $ant(b, a) > 0$. Les temps valides des observations et des sources étant flous, la détermination du sens temporel de a et b peut être complexe. Il est également possible qu'entre deux sources $s_1, s_2 \in S$ se chevauchant, certaines observations de s_1 précèdent celles de s_2 et que l'inverse se produise dans d'autres parties des sources.

Prenons par exemple le cas des deux observations a et b et de leur temps valides \tilde{t}_1 et \tilde{t}_2 illustrés dans la figure 5.6. Leurs fonctions d'appartenances sont respectivement $\tilde{t}_1 = (9, 10, 20, 21)$ et $\tilde{t}_2 = (13, 15, 40, 42)$. Dans l'intervalle $[9, 13]$, seule a peut être observée. Entre $t = 13$ et $t = 21$, les deux observations peuvent être faites. Enfin, entre $t = 21$ et $t = 42$, seule b peut être observée.

Dans l'intervalle $[13, 21]$, rien ne permet de savoir, sans connaissance supplémentaire, quelle observation a été faite en premier : il est possible que a précède b tout comme l'inverse. On a ainsi $ant(a, b) > 0$ et $ant(b, a) > 0$ qui traduisent cette situation. Il est donc possible qu'entre a et b , le sens des relations soient $b \rho_f a$. Les deux ordres décrits dans la figure sont donc valides, la relation entre b et a dans l'ordre 2 est dans le **sens rétro-temporel**.

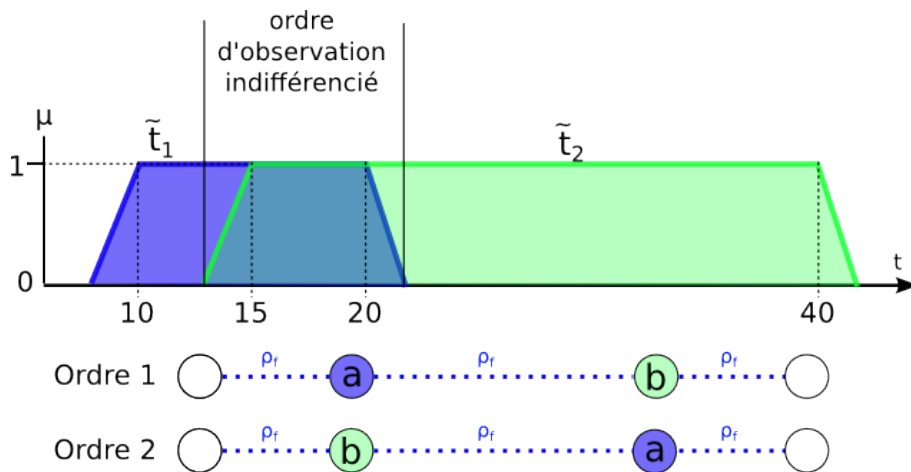


FIGURE 5.6 – Sens rétro-temporel : deux ordres possibles

Puisqu'il n'y a pas d'ordre total dans le cas des temps valides flous, nous devons relâcher toutes les contraintes fixant une succession entre instants dans le graphe

7. Rappelons le, le temps valide d'une observation est en réalité une période pendant laquelle il est possible que la trace de l'entité du monde réel ait été enregistrée dans une source.

spatio-temporel (déjà visible dans la définition des relations). Ce sens est appelé **local** car nous considérons qu'il peut s'appliquer uniquement à quelques relations entre deux sources, et non pas forcément aux deux sources entières⁸.

Continuations multiples : le cas des sources à différents niveaux de détail

Les sources étant pour nous des cartes, elles peuvent se présenter sous diverses formes et en particulier, à des échelles et des niveaux de détails différents. Prenons le cas d'une parcelle cadastrale représentée dans trois sources géohistoriques s_1, s_2, s_3 , illustré en figure 5.7. La première source présente une parcelle a , représentée également dans la seconde source (a' et a'') mais cette fois à un niveau de détail plus élevé qui fait apparaître une petite voie traversante. Enfin, la dernière source montre que deux parties de a'' lui ont été retirées pour former deux nouvelles parcelles b et c .

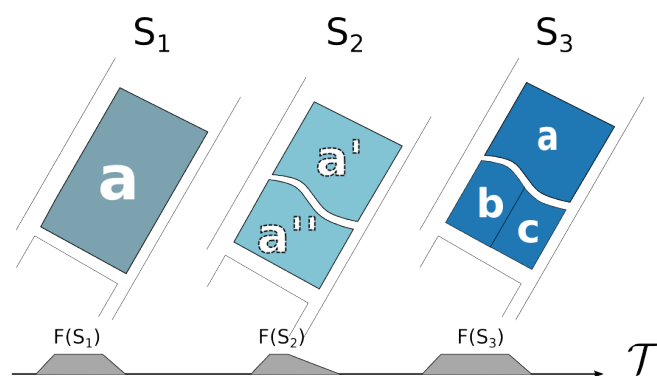


FIGURE 5.7 – Exemple de sources décrivant des observations à différents niveaux de détail.

L'observation a est exactement la même dans s_1 et s_2 , puis perd deux portions de sa surface dans la sources s_3 . Le graphe géohistorique résultant devrait donc être celui visible en figure 5.8. Or, toute la partie entourée en rouge est invalide dans le modèle de graphe géohistorique actuel. En effet, il viole les contraintes sur la transmission de l'identité (voir le paragraphe 5.1.3). En résultent deux observations de même identité dans une sources commune, ce qui est interdit par le graphe géohistorique. De plus, la suite des transformations est incohérente : a' continue dans s_3 tandis que a'' est divisée. La question du niveau de détail a été abordé par Del Mondo (2011) qui propose deux solutions : utiliser des opérateurs d'amalgamation ou faire évoluer le graphe spatio-temporel vers un bigraphe (voir Stell *et al.* (2011) et Milner (2008) de plus amples détails sur les bigraphes). L'utilisation de bigraphes pose un problème dans notre cas : nous ne connaissons pas par avance ni les différents niveaux d'abstraction des sources ni les règles de passage d'un niveau à l'autre. Autrement dit, nous sommes dans une situation où l'on cherche plutôt à découvrir ces niveaux

8. On pourrait aussi considérer que l'ordre temporel joue au niveau des sources uniquement. Ainsi, le sens rétro-temporel s'applique à toutes les observations liées de ces deux sources. Toutefois, rester au niveau des observations permet de conserver un plus grand niveau de généralité. De plus, le fait de rester au niveau des observations permet de prendre en compte l'hétérogénéité des périodes de levé des plans, où certaines parties peuvent être dans le sens de la flèche du temps, et d'autres inversées. Ce cas se présente en particulier lorsque les sources sont créées aux même périodes et que l'espace sous-jacent se transforme rapidement.

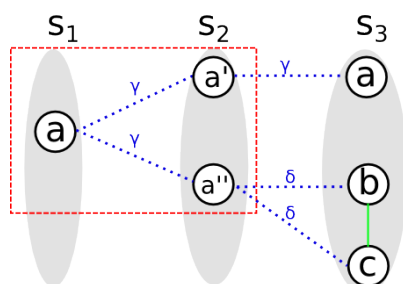


FIGURE 5.8 – Graphe géohistorique invalide résultant d’une différence de niveau de détail.

de détail et non les modéliser.

Si l’on considère que le cas de figure illustré en figure 5.7 se présente uniquement dans le cas de niveaux de détails hétérogènes, on peut amalgamer les observations a' , a'' de la source s_2 vers le niveau de détail le plus grossier. Concrètement, cela signifie que l’on peut, dans ce cas précis, créer un graphe géohistorique de niveau de détail homogène (aligné sur le plus faible) en agrégeant les nœuds de même identité dans une source. La figure 5.9 donne en exemple le graphe de la figure 5.8 une fois le niveau de détail réduit. Cette amalgamation doit être appliquée de façon récursive sur un graphe géohistorique pour résoudre tous les cas de continuations multiples.

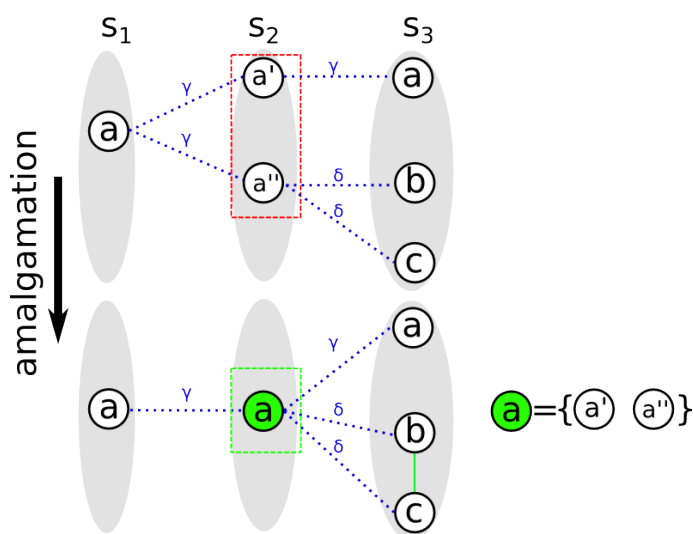


FIGURE 5.9 – Amalgamation dans le cas d’une continuation multiple.

Enrichissement de l’identité

Dans le graphe spatio-temporel, une entité a ne peut être une descendante d’une entité b en même temps qu’être un enfant d’une troisième entité c (voir la figure 5.10). Autrement dit, le processus d’annexion (un objet en absorbe un autre sans que cela change son identité) n’est pas représentable. Pour reprendre le cas des parcelles, cela signifie qu’il ne peut y avoir de parcelle ayant absorbé une part d’une parcelle adjacente. Or, ce type de situation existe couramment. Un cas est visible dans l’exemple

de graphe en figure 5.3 avec les parcelles a et b ⁹ Nous relâchons cette contrainte pour le graphe géohistorique.

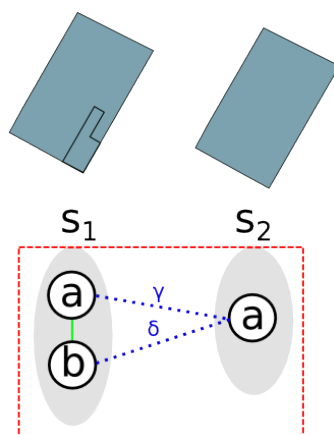


FIGURE 5.10 – Situation interdite par le graphe spatio-temporel, autorisée dans le graphe géohistorique.

5.2.4 Graphe géohistorique

Définition

L'ensemble des sources géohistoriques S , des observations σ et des relations qui existent entre ces dernières forment un **graphe géohistorique** $G_H(\sigma, E_H, S, \mathcal{T})$ dont σ est l'ensemble des noeuds et $E_H = \{(a, b) | a \rho b\}$ l'ensemble de ses arêtes. Tout comme le graphe spatio-temporel, le graphe géohistorique peut être décomposé en trois sous-graphes, suivant les types de relations auxquels correspondent ses arêtes. On a alors :

1. $G_s(\sigma, E_s, S, \mathcal{T})$ le graphe relatif aux relations spatiales, avec $E_s = \{(a, b) | a \rho_s b\} \forall a, b \in \sigma$
2. $G_{st}(\sigma, E_{st}, S, \mathcal{T})$ le graphe relatif aux relations spatio-temporelles, où E_{st} est défini de manière similaire à E_s .
3. $G_f(\sigma, E_f, S, \mathcal{T})$ le graphe relatif aux filiations, construit de la même manière que les deux autres.

La figure 5.11 présente un graphe géohistorique reprenant le même exemple que pour le graphe spatio-temporel. Cette fois, les relations de filiation des deux plus grandes parcelles entre Jacoubet et Vasserot sont des continuations.

5.2.5 Contraintes sur le graphe géohistorique

Certains choix de modélisation, en particulier l'organisation par source et le temps flou peuvent créer des graphes géohistoriques décrivant des transformations contre-intuitives ou brisant le principe de causalité. Nous posons ici deux contraintes sur le graphe géohistorique dans le but d'interdire la création de graphes invalides. Ces contraintes doivent être prises en compte au moment de l'instanciation du modèle dans une base de données.

9. Dont les évolutions avaient alors été considérées comme des dérivations.

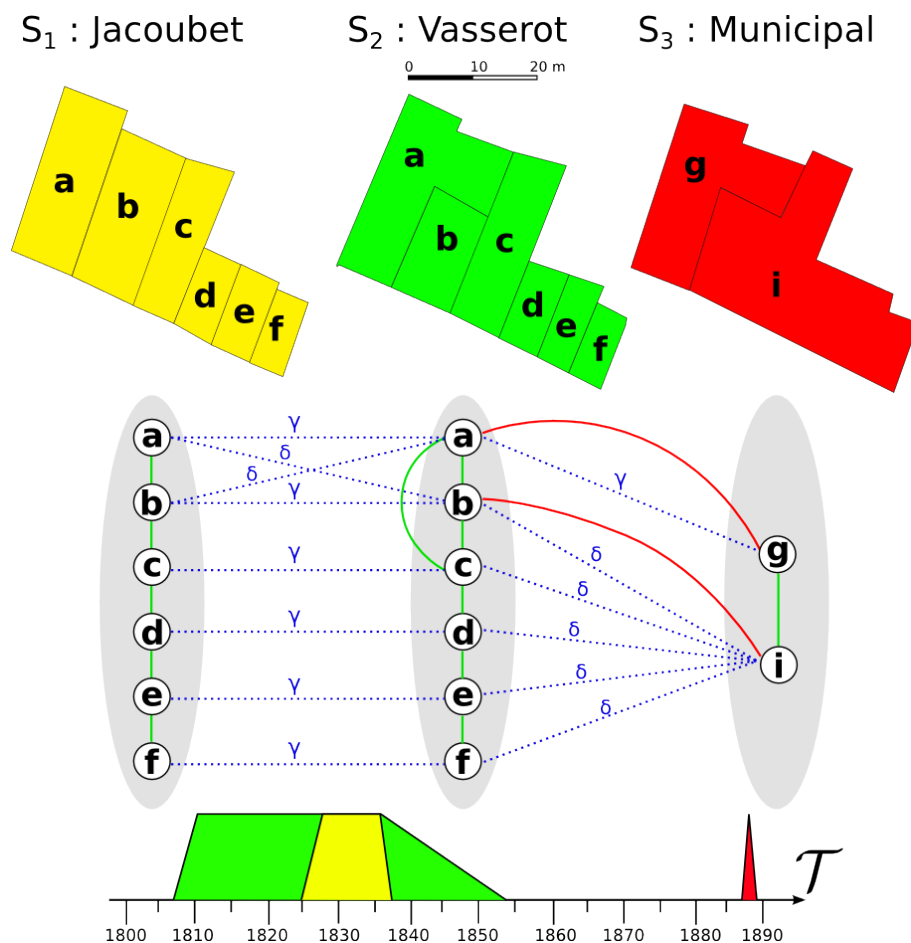


FIGURE 5.11 – Un exemple de graphe géohistorique.

Acyclicité

L'existence de relations de filiation rétro-temporelles rend possible l'apparition de cycles dans le graphe géohistorique. Non seulement ces cycles sont incompatibles avec une vision linéaire du temps, mais ils rompent également le principe de causalité : une observation peut être son propre ancêtre.

Nous posons une contrainte sur le graphe interdisant les cycles sur les relations de filiation, quelle que soit leur longueur. Autrement dit, une observation a du graphe géohistorique ne peut être son propre voisin en filiation :

$$\forall a \in \sigma \forall d \in \mathbb{N}, a \notin \rho_f^{+/-d}(a) \quad (5.7)$$

Transformations multisources

Dans le graphe géohistorique, nous considérons que chaque source est une représentation de l'espace, même si elle est incomplète. Implicitement, ceci signifie que les sources sont mutuellement exclusives et qu'il en va de même pour leur contenu. Par conséquent, un processus spatio-temporel ne peut avoir lieu entre plus de deux sources. Par exemple, une observation a dans une source s_1 ne peut pas être en

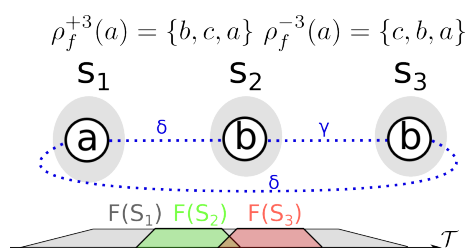


FIGURE 5.12 – Un exemple de graphe géohistorique cyclique.

relation de filiation avec une observation b dans s_2 et c dans s_3 (voir la figure 5.13). Tout comme l'acyclicité, cette contrainte doit être exprimée lors de l'instanciation. Cette contrainte implique que :

$$\begin{aligned} \forall a, b, c \in \sigma, a \rho_f b \wedge a \rho_f c &\implies s(b) = s(c) \\ \forall a, b, c \in \sigma, b \rho_f a \wedge c \rho_f a &\implies s(b) = s(c) \end{aligned} \quad (5.8)$$

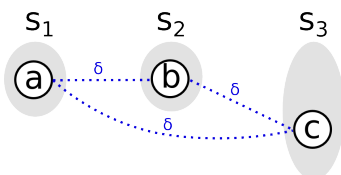


FIGURE 5.13 – Une transformation multisource.

5.3 Processus spatio-temporels

Le modèle de graphe spatio-temporel vise à extraire diverses connaissances sur les évolutions d'entités au cours du temps. Cependant, les processus spatio-temporels sont essentiellement utilisés comme contraintes d'intégrité lors de l'instanciation du modèle. Ceci signifie que le modèle de données permet de créer une base de données stockant les évolutions entre entités (de manière directe avec la dérivation, ou indirecte par contraintes) en assurant que ce stockage respecte certaines règles propres au phénomène représenté.

Notre problématique est différente : nous disposons déjà d'une base de données spatiale et temporelle très peu contrainte à partir de laquelle nous souhaitons extraire un graphe géohistorique. Nous sommes dans une démarche inductive où nous cherchons à détecter certains processus spatio-temporels (restreints aux transformations de l'espace) au sein des observations géohistoriques disponibles dans la base de données spatiale et temporelle.

Dans cette section, nous introduisons l'ensemble des types de processus spatio-temporels pertinents pour l'espace urbain. Nous nous appuyons pour cela sur la typologie des processus spatio-temporels établie par Claramunt et Theriault (1996) Cette typologie organise les processus en trois grands groupes :

- le premier concerne les évolutions individuelles des entités (pour nous observation) et en distingue deux types : les changements simples (stabilité, mort, naissance) et les transformations (déformation, déplacement, etc.),

- le second contient les transformations impliquant plus d'une entité : division, fusion, réallocation,
- le dernier traite des relations spatio-temporelles entre entités (diffusion, remplacement, etc.).

Nous avons choisi de représenter un certain nombre de processus spatio-temporels issus des deux premiers groupes de cette typologie que nous considérons pertinents pour modéliser les transformations d'observations d'un espace urbain.

Nous ajoutons également quelques processus que l'on trouve dans d'autres travaux (Galton, 2004; Hornsby et Egenhofer, 2000; Stefani, 2010), pertinents pour notre cas et s'intégrant aisément dans la typologie présentée. Tous les processus retenus se trouvent résumés en figure 5.14. Étant extraits d'un graphe géohistorique, **ces processus en sont des sous-graphes que l'on notera G** . La sémantique des différents processus ainsi que les règles permettant de les extraire d'un graphe géohistorique sont présentées dans la suite de cette section.

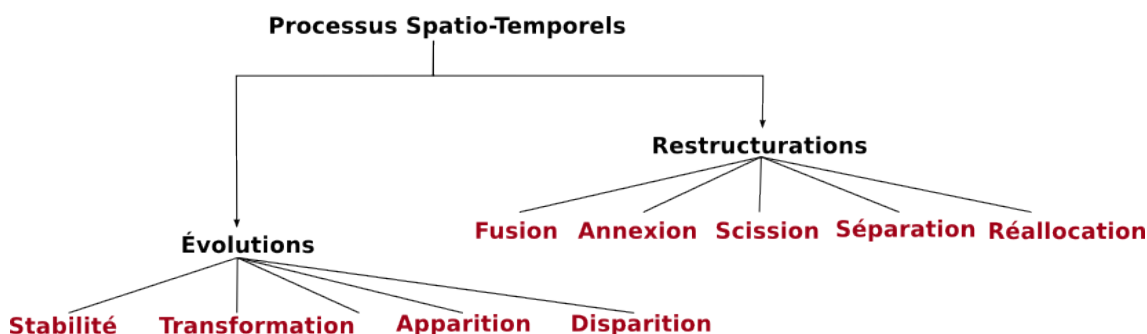


FIGURE 5.14 – Typologie des processus spatio-temporels représentés dans le graphe géohistorique

Évolutions : Stabilité et transformation

Dans la typologie d'origine, la stabilité ne concerne que la composante spatiale des entités, tandis que la transformation implique un changement de cette composante. Ainsi, deux observations liées sont stables même si leurs autres attributs peuvent changer. Ici, nous confondons les notions de stabilité et de continuation du graphe spatio-temporel. **Un processus de stabilité décrit donc l'existence d'une continuation entre deux observations.** De la même façon, **un processus de transformation correspond à une dérivation** dans le graphe géohistorique. Les figures 5.15 et 5.16 présentent graphiquement ces deux cas et donnent les règles permettant de les détecter dans un sous-graphe de filiation.

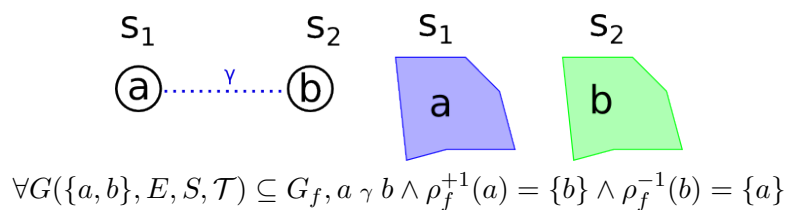


FIGURE 5.15 – Processus spatio-temporel de stabilité.

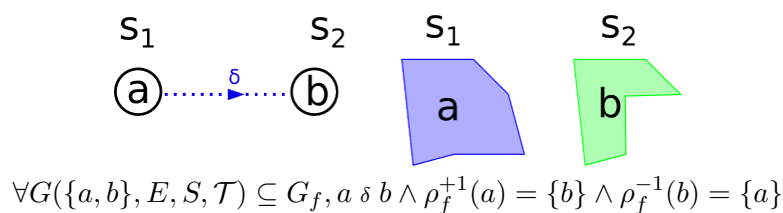


FIGURE 5.16 – Processus spatio-temporel de transformation.

Évolutions : Apparition et Disparition

Les processus d'apparition et de disparition décrivent respectivement la première et la dernière observation faite d'une même entité du monde réel. Cela ne signifie pas nécessairement que cette entité a disparu ou est apparue à ce moment là, mais qu'elle est visible dans une source géohistorique pour la première (respectivement dernière) fois à ce moment là. Ces deux processus sont décrits dans les figure 5.17 et 5.18.

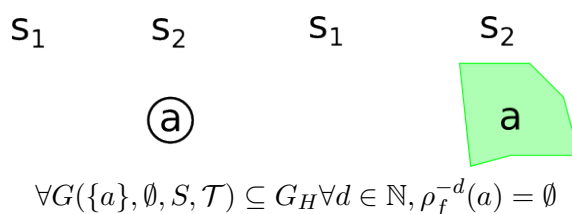


FIGURE 5.17 – Processus spatio-temporel d'apparition.

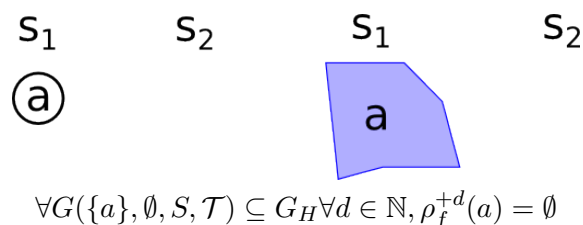


FIGURE 5.18 – Processus spatio-temporel de disparition.

Restructurations : Fusion / Scission

Il y a fusion lorsque plusieurs observations s'assemblent pour en former une nouvelle. C'est par exemple le cas de plusieurs parcelles cadastrales parties d'un même lot, vendues puis assemblées pour former une unique parcelle. Le processus de fusion est explicité par la figure 5.19. Le processus de scission est symétrique : une observation se trouve divisée en plusieurs autres (voir la figure 5.20).

Restructurations : Annexion et Séparation

Les processus d'annexion et de séparation sont des cas particuliers de fusion et de scission pour lesquels une des observations conserve son identité. Le processus d'annexion décrit le cas où une observation "absorbe" une autre observation sans pour autant que cela remette en question son identité. Si l'on reprend l'exemple

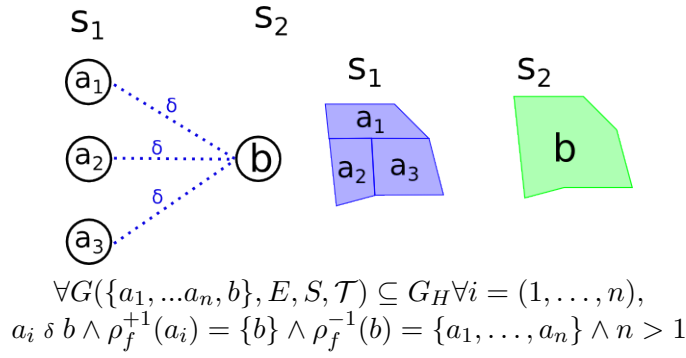


FIGURE 5.19 – Processus spatio-temporel de fusion.

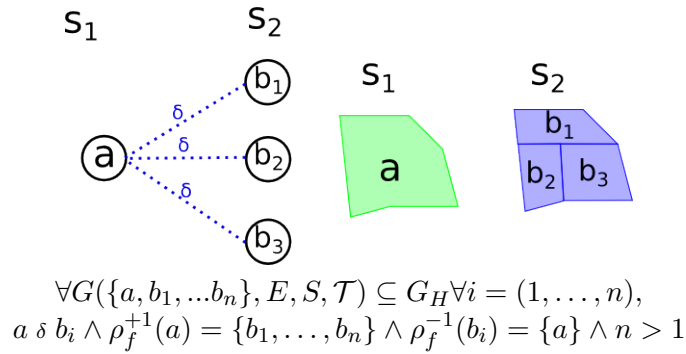


FIGURE 5.20 – Processus spatio-temporel de scission.

des parcelles, ce peut être le cas d’une parcelle s’étendant sur un fond de cour et intégrant une portion de cette cour. La séparation décrit le phénomène inverse : une observation perd une partie d’elle-même trop peu importante pour que cela change son identité. Les différences entre ces deux processus et ceux de fusion/scission sont liées à une définition de l’identité et donc au cas d’application ¹⁰.

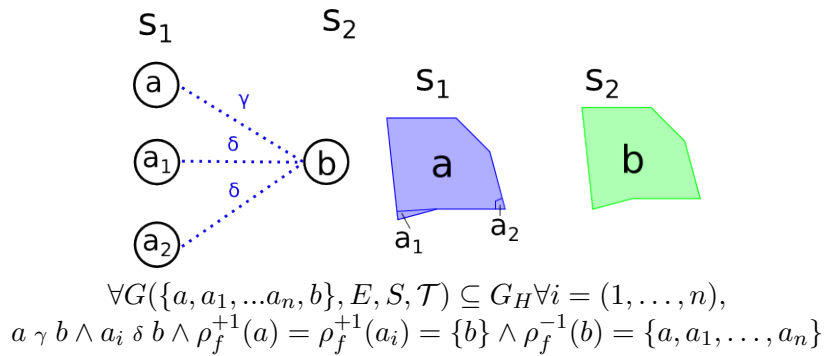


FIGURE 5.21 – Processus spatio-temporel d’annexion.

Restructurations : Réallocation

La réallocation est un processus particulier car il décrit la réorganisation d’un

10. Enlever une planche du bateau de Thésée est une séparation ou une scission ?

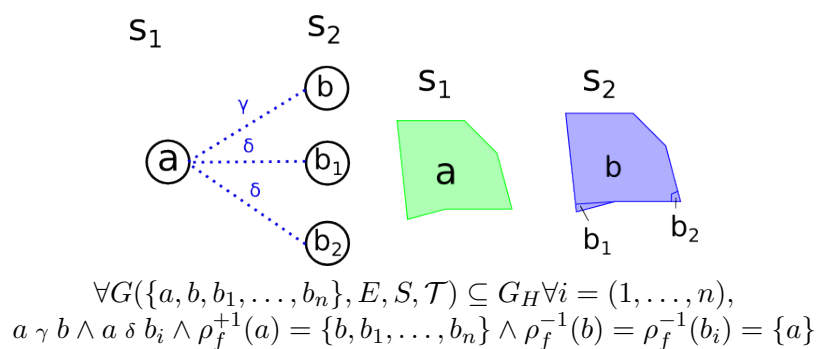


FIGURE 5.22 – Processus spatio-temporel de séparation.

ensemble d'observations en un autre ensemble. Un tel processus peut apparaître pour plusieurs raisons :

- la granularité temporelle est trop faible : certaines observations intermédiaires sont manquantes, ce qui empêche de décomposer la réallocation en une suite de processus plus précis (scission, fusion, etc.),
- la granularité spatiale est trop faible : les observations ne sont pas assez détaillées pour qu'il soit possible d'identifier précisément le phénomène sous-jacent,
- il s'agit réellement d'une réallocation (redistribution de parcelles en nouvelles parcelles lors d'une vente importante, etc.).

Quelle que soit la situation expliquant sa présence, nous considérons la réallocation comme un processus spatio-temporel à part entière. L'exemple fourni en figure 5.23 présente un cas ne contenant que des dérivations. Il est cependant tout à fait possible qu'une réallocation contienne également des continuations dès lors que celles-ci respectent les contraintes posées sur le graphe géohistorique.

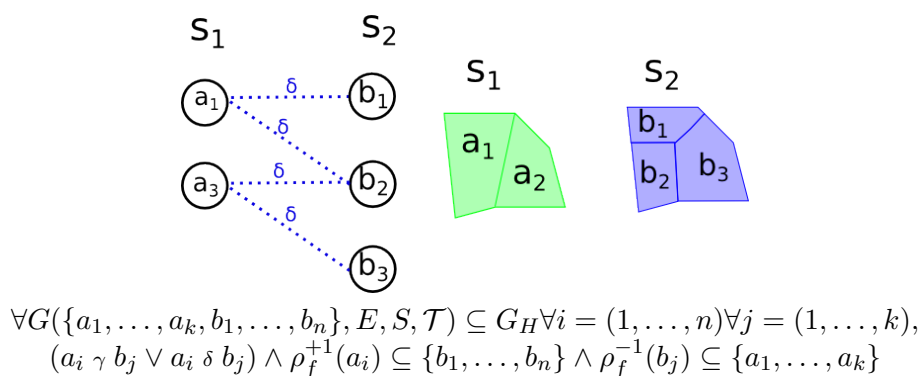


FIGURE 5.23 – Processus spatio-temporel de réallocation.

Processus spatio-temporels élémentaires

Tous les processus que nous avons définis dans cette section sont spécifiés de façon à être mutuellement exclusifs. Chaque observation ne peut donc être impliquée que dans deux processus spatio-temporels¹¹. Il ne peut donc y avoir de hiérarchie

11. L'un aboutissant à cette observation et le second en partant.

de différents processus pour un même ensemble d'observations. Par exemple, une réallocation ne peut pas être décomposée en fusion, scissions, annexions, etc. Nous considérons en effet que tous les processus posés sont **élémentaires**. Ce choix est motivé par le fait que ces processus décrivent des transformations des structures urbaines avec le niveau de détail le plus fin pouvant être atteint à partir d'observations géohistoriques fondamentalement imparfaites. Tenter de détailler plus des phénomènes qui ne nous sont accessibles que de façon grossière augmente le risque de sur-interprétation. Il s'agit donc d'une règle de prudence sur l'utilisation des connaissances issues des observations géohistoriques.

5.4 Conclusion

L'objectif de ce chapitre était de proposer une approche permettant d'extraire des transformations de l'espace urbain à partir d'une base de données spatiale et temporelle. Pour ce faire, il est nécessaire de s'appuyer sur un modèle de représentation de ces transformations. Nous avons choisi de nous appuyer sur un modèle spatio-temporel existant, proposé par Del Mondo (2011) qui permet de représenter les évolutions d'entités au sein d'une structure de graphe. Ce modèle, nommé *graphe spatio-temporel* permet d'inférer un certain nombre de connaissances sur les évolutions des entités par des fonctions de parcours dans le graphe. Ce modèle est également instancié au sein d'une base de données relationnelle.

Nous avons montré que l'approche de graphe spatio-temporel pouvait être adapté au cas des observations géohistoriques. Pour ce faire, nous avons modifié la proposition initiale afin, notamment, d'intégrer les temporalités floues des observations. Le modèle de graphe spatio-temporel ainsi modifié est nommé **graphe géohistorique**. Nous avons par la suite choisi neuf processus spatio-temporels pertinents pour la modélisation des transformations de l'espace urbain et nous avons proposé des règles permettant de les extraire d'un graphe géohistorique.

Afin qu'un graphe géohistorique puisse être extrait d'une base de données spatiale et temporelle, il faut connaître l'identité des différentes observations qui la compose et les relations de filiations entre ces observations. Dans le modèle de graphe spatio-temporel, l'identité est considérée connue par avance et intégrée explicitement dans la base de données au moyen d'identifiants. Notre cas est toutefois très différent sur ce point. En effet, nous ne connaissons pas par avance ni l'identité des observations, ni les relations de filiation qu'elles entretiennent entre elles. Il faut donc disposer d'un mécanisme permettant d'extraire ces identités et d'identifier les liens de filiation entre observations.

Chapitre 6

Construction du graphe géohistorique

Objectifs :

- Proposer une méthode permettant de construire de façon semi-automatique un graphe géohistorique à partir d'un ensemble d'observations vectorielles localisées dans le temps.
- Implémenter et expérimenter la méthode sur les réseaux viaires de Paris extraits par vectorisation des atlas parisiens.

Sommaire

6.1	Approche générale	240
6.1.1	Schéma général	240
6.1.2	Notions fondamentales	242
6.2	Découverte des relations de filiation	248
6.2.1	Un problème d'appariement de données géographiques multicritère	248
6.2.2	Éléments de la théorie des croyances	255
6.2.3	Approche par optimisation discrète	258
6.2.4	Un recuit simulé mono-objectif pour découvrir H_f^*	263
6.3	Typage des relations de filiations	277
6.3.1	Décomposition d'un hypergraphe H_f	278
6.3.2	Étiquetage des filiations par fusion d'informations incertaines . .	278
6.3.3	Construction d'un graphe géohistorique valide	280
6.4	Mise en œuvre et résultats : application aux rues de Paris . .	281
6.4.1	Implémentation de l'approche de construction de graphes géohis- toriques	281
6.4.2	Données utilisées	282
6.4.3	Fonctions d'agrégation des membres de l'identité des tronçons . .	286
6.4.4	Mesures de similarité d'identité utilisées pour la découverte des relations de filiation	288
6.4.5	Fonctions de masse de croyance utilisées pour le typage des filiations	291
6.4.6	Résultats sur l'îlot de la Trinité	292
6.4.7	Résultats sur le quartier des Arts et Métiers	299
6.5	Conclusion	306

Nous disposons une base de données spatiale et temporelle stockant des observations géohistoriques sous forme vecteur, temporalisées et organisées par source, au dessus de laquelle est défini un modèle de graphe géohistorique permettant de représenter les transformations entre ces différentes observations. Ce graphe, issu d'un modèle de graphe spatio-temporel de la littérature, doit être instancié à partir des différentes observations. Nous avons expliqué dans le chapitre 5 que cette instanciation impliquait de connaître par avance les liens de filiation existant entre les différentes observations stockées. Or, dans notre cas, ceux-ci ne sont pas connus a priori et doivent être identifiés. Dans ce but, nous proposons dans ce chapitre une approche permettant de retrouver ces liens de filiation, ce qui permet ensuite d'instancier un graphe géohistorique. Cette approche, présentée de façon globale en section 6.1, se compose de deux étapes.

Tout d'abord, les observations dont les identités sont liées sont découvertes par un processus d'appariement de données géographiques, ce qui produit un ensemble de liens de filiations. L'approche est présentée en détails en section 6.2.

Dans un second temps (cf. section 6.3) , le type de ces liens de filiation -dérivation ou continuation- est précisé à la volée à partir d'une définition de l'identité par l'utilisateur. Ainsi, cela nous permet de produire une instanciation du graphe géohistorique tel que nous l'avons défini dans le chapitre précédent.

L'approche complète est appliquée aux réseaux viaires de Paris en section 6.4 dans le but de reconstituer le graphe géohistorique retraçant les transformations des rues de la ville. Cette phase nous permet de valider l'approche et ainsi de finaliser l'ensemble des étapes permettant de créer une base de données spatio-temporelle sur l'espace urbain à partir de sources cartographiques anciennes.

6.1 Approche générale

Le graphe spatio-temporel sur lequel s'appuie notre graphe géohistorique est une structure inférée à partir d'une base de données stockant des entités, leurs identités et les relations de dérivation entre entités. Dans notre cas, l'identité des observations n'est pas définie au moment de la création de la base de données spatiale et temporelle dans la mesure où cette identité dépend de la nature de l'observation et du point de vue de l'utilisateur qui conçoit le graphe géohistorique. Par conséquent, les relations de filiation (dérivation ou continuation) ne sont pas connues a priori et il n'est pas possible d'inférer directement un graphe géohistorique à partir d'une base de données spatiale et temporelle. Il est toutefois nécessaire de connaître ces relations de filiation pour détecter les processus spatio-temporels existants entre observations issues de diverses sources géohistoriques. **Afin de retrouver les relations de filiation, il faut détecter les observations dont les identités sont liées.** Pour atteindre cet objectif, nous avons besoin d'une façon d'exprimer l'identité de nos observations géohistoriques, ainsi que d'une approche permettant, à partir d'une définition de l'identité, de détecter les relations de dérivation et continuation existant entre ces observations.

Nous présentons dans cette section les grandes étapes de l'approche que nous proposons pour atteindre ce but. Nous introduisons de plus quelques notions fondamentales sur lesquelles repose notre approche.

6.1.1 Schéma général

L'approche que nous proposons pour la construction d'un graphe géohistorique à partir d'un ensemble d'observations est résumée dans la figure 6.1. La présentation des deux étapes principales de l'approche qui apparaissent en rouge constitue le cœur de ce chapitre. Nous décrivons ici l'enchaînement général des étapes ainsi que les interactions avec l'utilisateur lors de la construction du graphe.

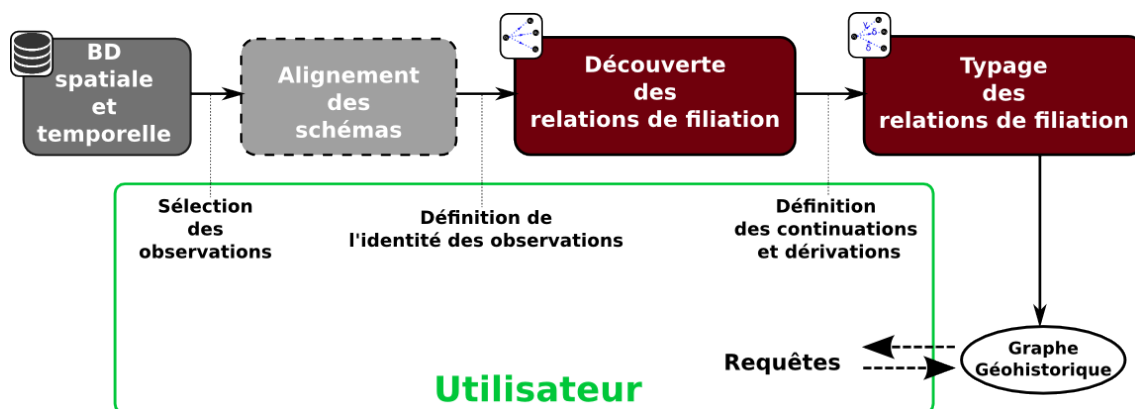


FIGURE 6.1 – Schéma général décrivant l'instanciation d'un graphe géohistorique à partir d'un ensemble d'observations.

Sélection des observations à lier et alignement des schémas

La base de données spatiale et temporelle stocke un ensemble d'observations géohistoriques pouvant être de nature variées. Ainsi, pour l'espace urbain, il peut s'agir d'observation de rues, de bâtiments, d'îlots, voire de données sociales géolocalisées (mouvements de foule, barricades, manifestations, etc.). Construire un graphe géohistorique entre des observations de nature aussi variée n'a pas toujours de sens ou n'est pas toujours pertinent pour une étude particulière¹. Cette étape de sélection correspond donc simplement à une sélection de données dans la base spatiale et temporelle.

Puisque la découverte de relations de filiation s'appuie sur la ressemblance entre les identités des observations qui composent le graphe, il faut que ces identités puissent être comparées. Celles-ci étant exprimées comme un ensemble d'attributs dépendant du *feature type* de l'observation, comparer deux identités signifie que les *feature types* des observations soient alignés. La plupart des approches proposées en géomatique pour aligner différents schémas de données géographiques reposent sur l'utilisation d'une ontologie pivot (Abadie, 2012, p.50), (Euzenat *et al.*, 2007). Cette ontologie support fournit un ensemble de concepts qui permettent de représenter les différents schémas et de déduire les relations de correspondance entre éléments de schémas à partir de ces annotations.

Cette ontologie dépend donc largement de la nature des observations stockées. L'appariement de schémas est un domaine à part. Dans le cadre de cette thèse, nous avons choisi de ne pas considérer l'alignement de schémas pour nous concentrer sur les étapes de construction du graphe géohistorique. En effet, étant donnée les données vectorisées, cet alignement a pu être effectué manuellement sans difficultés. Pour cette raison, nous considérons toujours que les *feature types* des observations sont alignés et donc que leurs identités sont comparables.

À la fin de cette étape, nous disposons d'un ensemble d'observations géohistoriques issues de sources différentes et dont les *feature types* sont alignés.

Découverte des relations de filiation

Cette étape constitue la première partie du processus permettant la découverte des relations de filiation entre les observations sélectionnées. La découverte des relations de filiation, présentée en détail dans la section 6.2, est un processus automatique permettant de détecter les observations dont les identités sont liées qui sont membres d'un des processus spatio-temporels introduits dans le chapitre 5.

Afin que la découverte des relations de filiation soit possible, il faut disposer d'une définition de l'identité des observations, c'est-à-dire d'un ensemble d'*attribute types* définissant quels attributs de chaque observation sont membres de son identité. La définition de l'identité dépend de la nature des observations mais également du point de vue de l'utilisateur créant le graphe géohistorique².

L'étape de découverte des relations de filiation permet de créer un premier graphe

1. Une rue n'est pas en relation de filiation avec une manifestation. Toutefois, on peut considérer qu'une rue peut être en relation de filiation avec un bâtiment (afin d'analyser le remembrement du bâti du aux percements par exemple).

2. Un utilisateur peut par exemple considérer que la forme et la localisation d'une rue constituent son identité, tandis qu'un autre peut également inclure le nom de la rue comme membre de l'identité. Dans le second cas, une rue changeant de nom impliquera un changement d'identité et donc la présence de deux rues affiliées. Dans le premier cas, il s'agit de la même rue qui continue à exister

de filiation dans lequel les observations ayant des identités liées sont identifiées.

Typage des relations de filiation

Une fois les observations liées identifiées, il faut déterminer si cette relation implique un changement d'identité ou si les identités des observations sont suffisamment proches pour que l'on considère qu'il s'agit de deux observations de la même entité du monde réel. Cela revient donc, pour chaque relation de filiation découverte, à la typer selon qu'il s'agit d'une continuation ou d'une dérivation.

Nous séparons cette étape de celle de la découverte des relations. En effet, les observations étant imparfaites, il peut être difficile de déterminer la différence entre continuation et dérivation³. De plus, la limite entre continuation et dérivation dépend ici encore de l'utilisateur⁴.

Le typage des relations de filiation issues de l'étape de découverte nécessite de disposer d'une définition fine de l'identité pour distinguer continuations et dérivations. En entrée de cette phase de typage, il faut donc que l'utilisateur précise sa définition de l'identité afin de pouvoir créer le graphe géohistorique final.

Une fois l'étape de typage terminée, nous disposons d'un graphe géohistorique complet sur l'ensemble des observations sélectionnées dans la base de données spatiale et temporelle.

6.1.2 Notions fondamentales

Formalisation d'une observation géohistorique

Pour la suite de ce chapitre, nous avons besoin de définir une observation géohistorique de façon formelle. Une observation est stockée dans la base de données spatiale et temporelle sous la forme d'un *feature* géographique, c'est-à-dire un n-uplet d'attributs dont les natures sont définies par un ensemble d'*attribute types*. Chaque observation est décrite par un *feature type* qui contient la liste des *attribute types* de l'observation.

Notons \mathcal{U} l'union des domaines de valeurs de tous les *attribute types* de la base de données et O un ensemble d'observations issues de la base de données spatiale et temporelle. Nous introduisons une fonction U_i qui associe à toute observation la valeur de l'attribut correspondant à son $i^{\text{ème}}$ *attribute type* : Une observation géohistorique

$$\begin{aligned} U_i : O &\rightarrow \mathcal{U} \\ U_i(o) &\mapsto u_i \end{aligned}$$

$o \in O$ peut donc être écrite

$$o = (U_1(o), U_2(o), \dots, U_n(o), U_g(o), U_t(o), U_s(o))$$

où :

3. Par exemple, les erreurs planimétriques créent des décalages qui gênent l'identification des continuations.

4. Est-ce que deux observations d'une rue ayant été élargie de quelques mètres doivent être en relation de dérivation ou de continuation ? Si l'utilisateur s'intéresse aux alignements des rues qu'il souhaite détecter et analyser, la relation doit être une dérivation. S'il s'intéresse aux transformations plus importantes, la relation peut être une continuation.

- $\{U_1(o), U_2(o), \dots, U_n(o)\}$ est un ensemble d'attributs thématiques quelconques pris dans \mathcal{U} ,
- $U_g(o)$ est sa représentation géométrique conforme au standard OGC *Simple Feature Access*⁵,
- $U_t(o)$ est le temps valide de l'observation représenté par un sous-ensemble flou,
- $U_s(o)$ est la source géohistorique de laquelle provient l'observation.

U_g, U_t et U_s correspondent respectivement à U_{n+1}, U_{n+2} et U_{n+3} . Nous les notons g, t, s pour des raisons de lisibilité.

Observations agrégées

Nous introduisons ici la possibilité de créer des observations agrégées à partir de plusieurs autres observations. Ces observations agrégées sont utilisées dans l'étape de découverte des relations de filiation. **Une observation agrégée est constituée de l'agrégation des attributs des observations qui la composent.** Par souci de simplicité, nous considérons que ces observations sont également représentables dans la base de données, c'est-à-dire qu'elles appartiennent également à O . Pour donner une définition de ces observations, nous devons étendre la définition de U_i à plusieurs observations. On pose : La fonction Ψ est une fonction d'agrégation

$$\begin{aligned} U_i : \bigcup_{k=0}^{\infty} O^k &\rightarrow \mathcal{U} \\ U_i(o_1, \dots, o_j) &\mapsto \Psi(U_i(o_1), U_i(o_2), \dots, U_i(o_j)) \end{aligned}$$

définie par : Le comportement de Ψ dépend du *feature attribute* correspondant à U_i .

$$\begin{aligned} \Psi : \bigcup_{k=0}^{\infty} U_i^k &\rightarrow \mathcal{U} \\ \{U_{i_1}, U_{i_2}, \dots, U_{i_j}\} &\mapsto u_i \end{aligned}$$

Par exemple, l'agrégation de géométries est une opération différente de l'agrégation de chaînes de caractères. Ces fonctions d'agrégation sont propres au type d'observation et aux choix de représentation adoptés pour les différents attributs de ces observations. Nous en conservons donc ici une notation générique.

Finalement, une observation agrégée $o_{p\dots q} \in O$ formée des observations $\{o_p, \dots, o_q\}$ telle que $\forall o, o' \in \{o_p, \dots, o_q\}, U_s(o) = U_s(o')$ est de la forme :

$$(U_1(o_{p\dots q}), \dots, U_n(o_{p\dots q}), U_g(o_{p\dots q}), U_t(o_{p\dots q}), U_s(o_{p\dots q}))$$

La figure 6.2 présente un exemple simple d'une observation agrégée définie sur deux observations de parcelles cadastrales possédant comme attribut thématique un numéro.

Identité d'une observation géohistorique

5. Qui décrit les primitives géométriques simples du type point, courbes et surfaces ainsi que les types géométriques composés de plusieurs primitives. Voir : <http://www.opengeospatial.org/standards/sfs>.

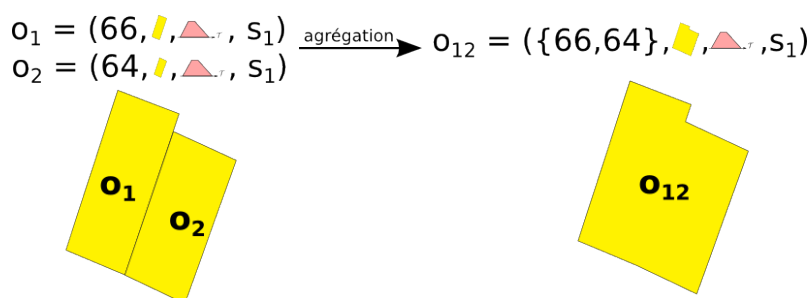


FIGURE 6.2 – Un exemple d'observation agrégée.

Nous avons déjà défini de façon générale l'identité d'une observation comme un ensemble de caractéristiques qui la distingue des autres. Ici, nous considérons que l'identité d'une observation géohistorique stockée comme une *feature* géographique est **un sous-ensemble de ses attributs thématiques et géométriques qui la distingue des autres observations issues de la même source géohistorique**. Cette identité peut être mise à profit pour retrouver au sein d'autres sources géohistoriques des observations présentant la même identité et étant donc d'autres représentations de la même entité du monde réel. Ainsi, le temps valide n'est pas un élément constitutif de l'identité. En effet, deux sources géohistoriques de même temps valide peuvent décrire deux observations de la même entité du monde réel. Nous pouvons maintenant préciser la fonction d'identité i posée dans le paragraphe 5.1.2 du chapitre 5.

L'identité d'une observation $o = (U_1(o), \dots, U_n(o), U_g(o), U_t(o), U_s(o)) \in O$, notée I_o , est définie par : où \mathcal{I} est l'ensemble des identités exprimables dans une base de

$$\begin{array}{l}
 i : O \rightarrow \mathcal{I} \\
 i(o) \mapsto I_o \subseteq (U_1(o), \dots, U_n(o), U_g(o))
 \end{array}$$

données spatiale et temporelle.

Hypergraphe de filiation

L'existence d'une relation de filiation entre deux observations n'a de sens que parce qu'elle décrit une partie ou la totalité d'un des processus spatio-temporels que nous avons définis. Découvrir des relations de filiations revient finalement à découvrir les processus spatio-temporels à l'œuvre entre les différentes observations. Pour ce faire, il faut disposer d'une structure modélisant ces processus spatio-temporels. Pour cela, nous introduisons au dessus d'un graphe de filiation $G_f(O, E_f, S, \mathcal{T})$ un **hypergraphe orienté** $H_f(O \cup o_\emptyset, E, S, \mathcal{T})$ nommé **hypergraphe de filiation** où :

- o_\emptyset est l'observation nulle, décrite dans la suite du paragraphe,
- E est un ensemble d'hyperarcs décrivant les processus spatio-temporels que l'on peut extraire de G_f ,
- S est l'ensemble des sources géohistoriques des observations de O ,
- \mathcal{T} est le domaine temporel flou.

Chaque type de processus spatio-temporel peut être décrit par un hyperarc. Un hyperarc $e \in E$ est un couple formé de deux sous-ensembles disjoints d'observations

TABLE 6.1 – Modélisation des processus spatio-temporels sous la forme d’hyperarcs

Graphe de filiation	Processus spatio-temporel	Hyperarc
S_1 S_2 	Apparition	
	Disparition	
	Continuation, Dérivation	
	Fusion, Annexion	
	Scission, Séparation	
	Réallocation	

géohistoriques (Q, T) . Le sous-ensemble Q est la **queue** de l’hyperarc, et T sa **tête**. Dans la suite de ce mémoire, nous écrirons $Q(e)$ et $T(e)$ les fonctions renvoyant respectivement la queue et la tête de l’hyperarc e .

Le tableau 6.1 présente la modélisation sous forme d’hyperarc de chaque processus spatio-temporel. Les cas des processus d’apparition et de disparition posent problème car ils correspondent à l’absence de relation de filiation. Pour qu’ils soient modélisables également au sein de H_f , nous introduisons une **observation nulle** $o_\emptyset = \emptyset$ virtuelle. Toute observations appartient ainsi à deux hyperarcs, l’un décrivant ce qui la précède, et l’autre ce qui lui succède. Il y a apparition ou disparition d’une observation si celle-ci est en relation de filiation avec cette observation nulle. Un exemple d’hypergraphe de filiation construit à partir du graphe de filiation G_f sur les parcelles cadastrales (voir le chapitre 5) est illustré en figure 6.3.

Hypergraphes et graphes de filiation admissibles

Nous avons posé quelques contraintes sur les processus spatio-temporels dans le chapitre 5. Nous devons redéfinir ces contraintes sur l’hypergraphe de filiation. Nous introduisons également une contrainte de Berge-acyclicité⁶ sur cet hypergraphe.

Un hypergraphe de filiation $H_f(O \cup o_\emptyset, E, S, \mathcal{T})$ est admissible s’il respecte les contraintes suivantes :

- C_1 : **Processus multisource**. Pour tout $e \in E$, toutes les observations de $Q(e)$ proviennent de la même source géohistorique S_i et toutes les observations de $T(e)$ provienne d’une seconde source géohistorique S_j avec $i \neq j$.

6. La notion de graphe acyclique sur un hypergraphe est plus complexe que pour un graphe simple et n’est pas unique. Il en existe 4 types, dont le détail peut être trouvé dans Duris *et al.* (2009).

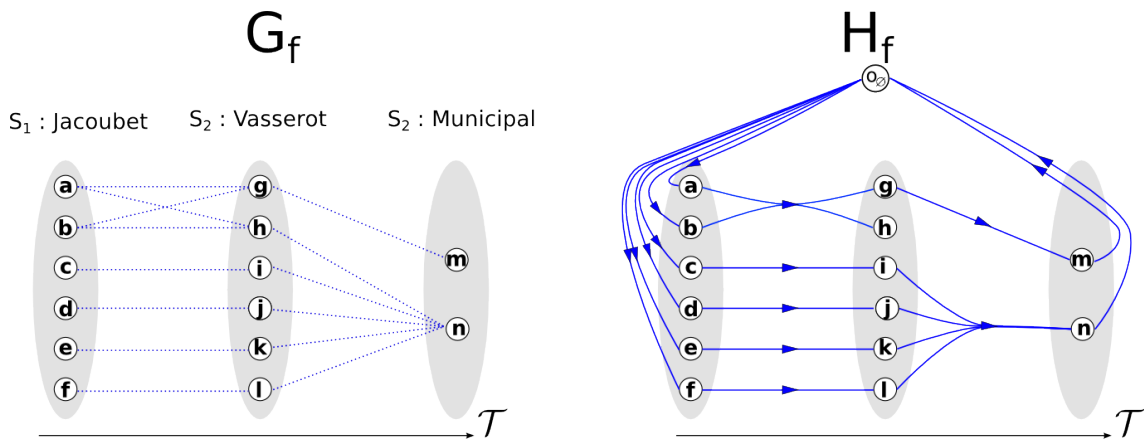


FIGURE 6.3 – Un hypergraphe de filiation représentant les processus spatio-temporels d’un graphe de filiation.

C_2 : **Relations rétro-temporelles.** Pour tout $e \in E$ avec o_Q l’observation agrégée formée des observations $Q(e)$ et o_T l’observation agrégée formée de $T(e)$, on a $ant(U_t(o_Q), U_t(o_T)) > 0$.

C_3 : **Berge-acyclicité.** Un Berge-cycle dans un hypergraphe de filiation est une suite $(e_1, o_1, \dots, e_n, o_n)$ telle que :

- . les hyperarcs e_i sont distincts,
- . les sommets o_i sont distincts,
- . pour tout i , o_i appartient à $T(e_i)$ et à $Q(e_{i+1})$

Un hypergraphe Berge-acyclique est un hypergraphe sans Berge-cycle.

Des exemples d’hypergraphes non admissibles sont visibles en figure 6.4. Notons que la contrainte C_1 empêche également la création de boucles dans l’hypergraphe.

Un graphe de filiation est admissible si l’hypergraphe de filiation que l’on peut construire à partir de ses relations est admissible.

Dans la suite du document, nous utiliserons par souci de simplicité le terme ”arc” pour désigner un hyperarc de H_f lorsqu’il n’y a pas ambiguïté.

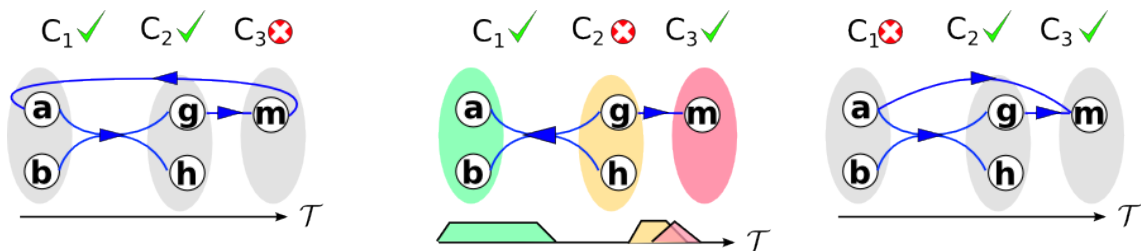


FIGURE 6.4 – Exemples d’hypergraphes de filiation non admissibles.

Fonctions de vraisemblance d’un hyperarc de filiation

Pour découvrir les relations de filiation au sein d’un ensemble d’observations, il faut être capable de déterminer si les processus spatio-temporels que l’on peut extraire à partir de ces relations correspondent à des **transformations vraisemblables des entités du monde réel sous-jacentes**.

Cela revient à juger de la vraisemblance d'un hyperarc dans H_f . Dans ce but, nous définissons une **fonction de vraisemblance d'un hyperarc de filiation**⁷ Φ définie pour tout hyperarc $e \in E$ d'un hypergraphe de filiation $H_f(O \cup o_\emptyset, E, S, \mathcal{T})$ par : avec :

$$\begin{aligned}\Phi : E &\rightarrow [0, 1] \\ \Phi(e) &\mapsto \phi(o_Q, o_T)\end{aligned}$$

- o_Q l'observation agrégée à partir des observations de $Q(e)$,
- o_T l'observation agrégée à partir des observations de $T(e)$,

Enfin, $\phi : O \times O \rightarrow [0, 1]$ est une fonction qui évalue la similarité des identités des observations agrégées formant la queue et la tête de l'hyperarc et la bonne succession temporelle de ces observations. La valeur de ϕ résulte donc d'une fusion de plusieurs mesures.

Mesures utilisées pour estimer ϕ

Pour pouvoir calculer ϕ , nous introduisons deux catégories de mesures entre deux observations o_1 et o_2 de O , simples ou agrégées :

- des **mesures de similarité d'identités** qui évaluent la ressemblance entre deux attributs de même *attribute type* au sein des identités de deux observations géohistoriques,
- des **mesures de succession** qui évaluent si o_2 succède à o_1 et lui est consécutive.

Une mesure entre deux observations o_1 et o_2 , notée c , est de la forme suivante : où :

$$\begin{aligned}c : \mathcal{U} \times \mathcal{U} &\rightarrow [0, 1] \\ c(U_i(o_1), U_i(o_2)) &\mapsto s\end{aligned}$$

- U_i est un *attribute type* commun à o_1 et o_2 ,
- s est la valeur mesurée normalisée sur $[0, 1]$.

Les mesures de similarité d'identité dépendent du type des attributs constituant l'identité des observations et ne peuvent donc être définies que lorsque ces types sont connus. Nous renvoyons à la section 6.4.4 pour la définition de telles mesures dans le cas des rues de Paris.

Les mesures de succession portent sur les temps valides des observations et sur l'ordre de leurs sources géohistoriques. Elles sont indépendantes de la nature de ces observations, ce qui nous permet de les définir dès à présent. Dans le chapitre 4, nous avons proposé une mesure d'antécédence entre deux temps flous et une méthode d'ordonnement de quantités floues. Nous proposons de construire deux mesures de succession à partir de ces dernières.

Mesures de succession

La première mesure de succession, notée c_{ant} , évalue pour deux observations o_1 et o_2 à quel point o_1 est antécédente à o_2 . Plus cette antécédence s'approche de 1

7. La vraisemblance telle que nous la définissons ici n'a pas de rapport direct avec la notion de vraisemblance statistique.

plus il y a de chances pour que la trace à laquelle correspond o_1 représente une entité antérieure à l'entité représentée par o_2 . Si ces observations sont en relation de filiation, cela signifie que o_1 est parente de o_2 . Nous réutilisons directement la mesure d'antécédence choisie dans le chapitre 4 pour définir cette mesure :

$$c_{ant}(U_t(o_1), U_t(o_2)) = ant(U_t(o_1), U_t(o_2))$$

La seconde mesure de succession c_{rank} entre o_1 et o_2 évalue la distance -en nombre de sources- qui sépare la source de o_1 de celle de o_2 . Cette mesure s'appuie sur l'indice de tri de quantités floues Y_3 proposée par Yager (1981) et que nous avons retenu dans le chapitre 4.

Cette mesure revient à calculer le nombre de "sauts" qu'il faut effectuer pour passer de la source géohistorique de o_1 à celle de o_2 . Elle est définie par :

$$c_{rank}(U_s(o_1), U_s(o_2)) = e^{(-\frac{K}{\epsilon})}$$

$$K = |Rang(U_s(o_1)) - Rang(U_s(o_2))|$$

où :

- $Rang(S_i) \in \mathbb{N}$ est le rang de la source S_i après rangement à l'aide de l'indice Y_3 des sources décrivant les observations de O ,
- ϵ est un paramètre de seuil permettant d'agir sur le nombre de "sauts" autorisés avant de considérer la relation de filiation comme non vraisemblable. Ce paramètre ne peut être déterminé que si l'on connaît une valeur limite entre vraisemblance et non vraisemblance. La détermination de cette valeur sera présentée dans le paragraphe 6.2.3

6.2 Découverte des relations de filiation

Au sein de cette section, nous présentons l'approche que nous avons mise en œuvre pour identifier les relations de filiation entre observations d'un même *feature-type* stockées au sein d'une base de données spatiale et temporelle. Cette approche s'appuie sur une méthode d'appariement multicritère par optimisation prenant en compte les imperfections des observations géohistoriques.

Dans un premier temps, nous justifions le choix de l'approche par appariement que nous étayons avec un état de l'art des principales méthode d'appariement de données géographiques. Cet état de l'art nous révèle les insuffisances de ces approches pour notre objectif mais il nous permet d'identifier certains fondements théoriques et choix de conception pouvant être mis à profit dans notre situation.

Nous proposons, dans un second temps, une méthode d'appariement de données géographiques imparfaites par optimisation combinatoire permettant de découvrir les relations de filiation entre des observations géohistoriques issues d'un ensemble de sources.

6.2.1 Un problème d'appariement de données géographiques multicritère

La découverte des relations de filiation repose sur l'hypothèse selon laquelle deux observations membres d'un processus spatio-temporel ont des identités similaires.

Cela signifie qu'entre deux sources géohistoriques S_1 et S_2 , on recherche les relations de filiation décrivant les processus spatio-temporels les plus vraisemblables entre les différentes observations issues de S_1 et S_2 .

Les observations pouvant être agrégées, cela revient à comparer les identités de chaque paire d'observations et à rechercher les observations dont les identités sont les plus proches. Les observations étant des objets géographiques représentés sous forme de *features*, le problème revient donc à comparer des objets géographiques pour identifier des objets homologues au sein de différentes sources de données. La mise en correspondance de plusieurs bases de données géographiques dans le but d'identifier des objets géographiques homologues fait l'objet de recherches dans le domaine de la géomatique depuis de nombreuses années. On parle alors d'**appariement de données géographiques**. Le but des méthodes d'appariement de données géographiques est d'identifier les objets communs dans deux bases de données, parfois à des niveaux de détail différents. Cet appariement peut être réalisé pour des objectifs divers : fusion d'informations provenant des deux bases de données, transfert d'attributs, propagation de mise à jour, versionnement, etc. L'utilisation de méthodes d'appariement pour la détection de processus spatio-temporels a déjà été proposée, par exemple par De Runz (2008) ou Gombosi *et al.* (2003). Nous faisons ici un état de l'art des principales méthodes d'appariement de données géographiques afin de discuter de leur adaptabilité à notre problème. Afin que cette discussion soit possible, nous posons d'abord les fonctionnalités que doit proposer une telle méthode pour être capable de découvrir des relations de filiation entre observations géohistoriques.

Fonctionnalités nécessaires pour la découverte de relations de filiation entre observations géohistoriques

Pour permettre de découvrir les relations de filiation entre observations, une méthode d'appariement de données géographique doit proposer les fonctionnalités suivantes :

1. **Critères multiples.** Le but est de comparer les observations selon plusieurs critères portant sur leurs identités et leurs temporalités, ce qui implique de disposer de plusieurs mesures permettant ces comparaisons. La méthode d'appariement doit donc être **multicritère**.
2. **Sources multiples.** Les observations composant le graphe de filiation que l'on cherche à construire sont issues d'un ensemble quelconque de sources géohistoriques. S'il est envisageable d'effectuer un appariement de proche en proche⁸(cf. notamment (Costes *et al.*, 2015)), cette manière de procéder ne prend pas en compte certaines complexités des données géohistoriques. D'abord, rappelons que nous autorisons les relations de filiation à effectuer des "sauts temporels" : ces relations peuvent exister entre deux observations S_1 et S_3 , même s'il existe une source S_2 dont le temps valide se situe entre ceux des deux autres sources (voir le paragraphe 5.2.3 du chapitre 5). De plus, nous autorisons également des relations de filiation "rétro-temporelles" (voir le paragraphe 5.2.3 du chapitre 5) qui ne pourraient alors pas être détectées.

8. En identifiant d'abord les relations de filiation entre les sources les plus anciennes, puis entre ce résultat et la source suivante, etc.

Nous devons donc disposer d'une méthode permettant d'apparier plusieurs sources de données à la fois.

3. **Gestion des imperfections.** Nos observations géohistoriques sont des objets géographiques imparfaits. Elles peuvent être incomplètes si certains attributs manquent, imprécises, erronées (par exemple un décalage planimétrique), etc. Il est donc crucial de disposer d'une méthode prenant en compte des données imparfaites.
4. **Cardinalité des appariements.** La majorité des méthodes d'appariement s'intéresse à identifier, pour chaque objet géographique d'une première base de données, son objet homologue dans une seconde base. Il s'agit donc de trouver uniquement des liens de cardinalité 1 : 1 entre objets homologues. Cependant, il est possible qu'un objet ait plusieurs objets homologues. Ce cas se produit principalement lorsque le niveau de détail des bases de données est différent. Dans le cas des observations géohistoriques, les situations de cardinalités supérieures au cas 1 : 1 sont nombreuses car nous recherchons explicitement des cas où des observations se divisent ou s'assemblent. De plus, nous devons également gérer le cas des apparitions et disparitions. Nous avons donc besoin d'une méthode capable de détecter des observations non appariées (apparitions et disparitions, de cardinalité 0 : 1 ou 1 : 0), des filiations simples (continuations/dérivations, de cardinalité 1 : 1) et les différents cas de filiations avec des observations agrégées (fusion/annexion : N : 1, scission/séparation : 1 : N, réallocation : N : M).
5. **Type de géométrie.** Souvent, les méthodes d'appariement de données géographiques sont adaptées à des objets géographiques modélisés avec un type de géométrie unique (points, polygones ou polygones par exemple). Or, les observations géohistoriques pouvant décrire n'importe quelle entité du monde réel, nous n'avons aucun a priori sur le type de géométrie qui permet de représenter leur forme et leur localisation.

État de l'art des méthodes d'appariement de données géographiques

Nous avons réalisé un état de l'art des principales méthodes d'appariement de données géographiques que nous avons confrontées aux fonctionnalités nécessaires définies dans le paragraphe précédent. Il existe de nombreuses méthodes d'appariement de données géographiques, souvent destinées à un cas d'application particulier. Nous avons conservé ici seulement 8 propositions, sélectionnées selon les trois critères suivants :

- il s'agit d'une méthode régulièrement citée dans des travaux liés à l'appariement de données géographiques,
- la méthode est adaptée à la découverte de transformations entre objets géographiques,
- certains auteurs ont proposés d'appliquer la méthode au cas spécifique de la découverte des transformations entre objets géographiques.

De plus, nous nous sommes restreint aux méthodes d'appariement de données géographiques vectorielles, nos observations étant des *features* vecteur. Les différentes méthodes sont présentées dans le tableau 6.2. Chaque colonne correspond à une fonctionnalité nécessaire pour le cas des observations géohistoriques.

Nous avons également ajouté une courte description du verrou scientifique visé par chaque proposition. De plus, la plupart des méthodes sont adaptées à des objets géographiques représentés par un certain type de géométrie (point, polyligne, polygone, etc.). Enfin, la colonne "Type d'approche" indique si la méthode s'appuie sur un fondement théorique spécifique. L'approche est notée comme "algorithmique" lorsqu'elle ne s'appuie pas sur une théorie mathématique particulière. Il s'agit généralement de méthode fondée sur la comparaison des géométries des objets traités ou leurs relations topologiques.

Bilan et discussion sur les méthodes d'appariement de l'état de l'art

À la lecture du tableau 6.2, on peut constater qu'aucune méthode ne fournit toutes les fonctionnalités requises. Pour la majorité des méthodes, cela est dû au fait qu'elles sont mises en place pour un problème spécifique, ce qui réduit fortement la généralité de la méthode. Ainsi, les méthodes proposées par Mustière et Devogele (2008), Walter et Fritsch (1999b), Li et Goodchild (2011b) et Volz (2006) s'intéressent au cas spécifique de l'appariement de réseaux de voies. De plus, toutes les méthodes à l'exception de Samal *et al.* (2004) et Li et Goodchild (2011b) et surtout Olteanu (2008) s'appuient uniquement sur la géométrie (et éventuellement les relations topologiques) des objets géographiques pour identifier les objets homologues. Un tel choix s'explique par le fait que la forme et la localisation des objets géographiques en constituent les attributs les plus significatifs et, souvent, les seuls fiables ou les plus renseignés dans les bases de données utilisées par les auteurs (ou, tout du moins, pour leur définition de l'identité)

Parmi les méthodes présentées dans le tableau, les propositions de Olteanu (2008) et Samal *et al.* (2004) sont toutefois intéressantes car elles répondent toutes deux à quatre des cinq fonctionnalités dont nous avons besoin. Nous présentons ici la méthode de Samal *et al.* (2004) avant de nous attarder plus longuement sur la proposition d' Olteanu (2008). Pour appairier plusieurs ensembles d'objets géographiques Samal *et al.* (2004) proposent de construire un graphe complet⁹ entre tous les objets géographiques des différents ensembles, puis de ne conserver que les arêtes ne liant pas deux objets du même ensemble et de longueur inférieure à un seuil donné. Chaque arête est ensuite pondérée par une similarité calculée entre les deux objets géographiques liés. Cette similarité est une agrégation de plusieurs similarités s'exprimant sur divers attributs des objets géographiques, notamment la forme de leur géométrie, leur position, ainsi que des similarités entre attributs thématiques représentés sous forme de chaînes de caractères (leur nom, leur fonction, etc.). Le problème revient alors à identifier des groupes d'objets issus des différents ensembles reliés par des arêtes de poids maximal. Un aspect particulièrement intéressant de cette méthode est qu'elle intègre, parmi ses mesures de similarité, une notion de contexte local des objets géographiques : deux objets géographiques sont d'autant plus similaires que les objet les plus significatifs¹⁰ les entourant sont répartis de façon similaire. La méthode présente cependant deux faiblesses :

9. Un graphe complet est un graphe dont tous les sommets sont reliés deux à deux par une arête.

10. Il s'agit des objets les plus précis, ou les plus importants. Dans le cas de la mise en œuvre de cette méthode pour appairier des objets représentés par des polygones, il s'agit des n objets de plus grande surface.

TABLE 6.2 – Tableau résumant l'adéquation des principales approches d'appariement de données géographiques aux fonctionnalités nécessaires pour la découverte de relations de filiations.

Méthode	Multicritère	Sources > 2	Imperfections	Cardinalités
(Mustière et Devogele, 2008)	non	non	non	0:1, 1:0, N:1, 1:N, N:M
(Olteanu, 2008)	oui	non	oui	0:1, 1:0, N:1, 1:N, (N:M)
(Walter et Fritsch, 1999a)	non	non	décalages planimétriques	0:1, 1:0, N:1, 1:N, N:M
(Samal <i>et al.</i> , 2004)	oui	oui	oui	0:1, 1:0, 1:1
(Li et Goodchild, 2011b)	non	non	décalages planimétriques	0:1, 1:0, N:1, 1:N
(Gombosi <i>et al.</i> , 2003)	non	non	non	0:1, 1:0, 1:1
(Beeri <i>et al.</i> , 2004)	non	non	non	0:1, 1:0, 1:1, 1:N
(Volz, 2006)	non	non	non	1:1, 1:2

TABLE 6.3 – Tableau résumant l'adéquation des principales approches d'appariement de données géographiques aux fonctionnalités nécessaires pour la découverte de relations de filiation.

Méthode	Géométries	Type d'approche	Verrou
(Mustière et Devoegele, 2008)	lignes polygonales	algorithmique	Apparier deux réseaux routiers représentés à des niveaux de détail différents.
(Olteanu, 2008)	tout type	théorie des croyances	Apparier deux ensembles de données à partir de connaissances imparfaites.
(Walter et Fritsch, 1999a)	lignes polygonales	théorie de l'information	Apparier deux réseaux routiers et évaluer automatiquement la qualité de l'appariement.
(Samal <i>et al.</i> , 2004)	tout type	algorithmique	Apparier $N \geq 2$ ensembles de données simultanément.
(Li et Goodchild, 2011b)	lignes polygonales	optimisation combinatoire	Apparier deux réseaux de rues et corriger dynamiquement les décalages planimétriques entre eux.
(Gombosi <i>et al.</i> , 2003)	polygones	algorithmique	Détecter les changements entre deux ensembles de polygones.
(Beerli <i>et al.</i> , 2004)	points	algorithmique	Apparier efficacement deux ensembles d'objets géographiques ponctuels lorsque les ensembles se recouvrent peu ^a .
(Volz, 2006)	lignes polygonales	algorithmique	Apparier deux réseaux de rues rapidement et sans perte de qualité par rapport aux méthodes existantes.

^a. Il y a faible recouvrement entre deux ensembles de données lorsque très peu d'objets du premier ensemble possèdent un homologue dans le second ensemble.

- d'une part, le calcul est extrêmement coûteux car toutes les combinaisons d'objets géographiques en dessous du seuil de distance maximal doivent être évaluées, même si une seule combinaison est finalement conservée,
- d'autre part, cette méthode ne permet d'identifier que des appariements de cardinalité maximale 1 : 1.

Ces limitations -en particulier sur les cardinalités- nous obligent à éliminer cette méthode. En effet, la possibilité de détecter des appariements de cardinalités supérieures à 1 : 1 est au cœur de notre problème. La proposition de Samal *et al.* (2004) pourrait être étendu au cas 1 : N si l'on ajoute une étape supplémentaire consistant à agréger des objets géographiques d'une même base de données lorsque ceux-ci sont fortement similaires à un même objets d'une autre base. De cette façon, l'approche n'est pas modifiée en profondeur. Il n'est toutefois pas possible de prendre en compte de la même façon les cas N : 1 et N : M.

La proposition d'Olteanu (2008) s'appuie sur la théorie des croyances, ou théorie de l'évidence (de l'anglais *theory of evidence*), introduite par Dempster (1968) et formalisée par Shafer (1976). La théorie des croyances permet de modéliser des propositions (hypothèses, évènements, décisions) sur lesquelles des sources d'informations imparfaites se prononcent. Cette théorie permet en particulier de représenter des informations incertaines ou imprécises¹¹. Nous présentons ici l'approche d'Olteanu de façon très générale. Pour plus de précisions, le lecteur peut se rapporter à la thèse complète citée dans la bibliographie du présent mémoire.

L'approche d'Olteanu (2008) appliquant cette théorie dans le cadre de l'appariement de données géographiques peut se résumer de la façon suivante :

1. Un ensemble de critères mesurent la similarité entre les différents attributs des objets géographiques issus de deux jeux de données (l'un *source* et l'autre *cible*). Ces critères peuvent porter sur la géométrie, la topologie ou les attributs thématiques des objets.
2. Pour chaque objet géographique du premier jeu de données, un ensemble d'objets "candidats homologues" du second jeu de données est sélectionné selon une heuristique fondée sur une distance maximale. Sur chaque couple formé de l'objet à apparier et d'un candidat, chaque critère exprime un degré de croyance en l'une des trois hypothèses : que les objets soient homologues, qu'ils ne le soient pas ou que la situation est totalement incertaine.
3. Pour chaque couple, une croyance globale pour chaque hypothèse est ensuite calculée en fusionnant les différentes croyances issues des critères de similarité.
4. Enfin, le candidat ayant la plus grande croyance dans le fait qu'il soit l'homologue de l'objet à apparier est conservé comme meilleur candidat. Si aucun candidat n'est un homologue crédible, l'objet est jugé non apparié. Les appariements de cardinalité 0 : 1, 1 : 0 et 1 : 1 sont détectés de cette façon.
5. Les appariement de cardinalité N : 1 sont détectés en regroupant a posteriori les appariements de cardinalité 1 : 1 lorsque un candidat a été choisi par

11. Nous ne présentons pas plus en détail la théorie des croyances dans ce paragraphe. Elle sera abordée plus avant dans la suite du mémoire.

plusieurs objets du premier jeu de données. Les appariements de cardinalité $1 : N$ sont détectés en reproduisant le même processus mais en inversant les deux jeux de données (le jeu *source* devient *cible* et vice-versa). Enfin, les appariements de cardinalité $N : M$ peuvent être trouvés en regroupant les appariements $1 : N$ et $N : 1$.

La méthode d'Olteanu (2008) présente l'avantage majeur de modéliser des connaissances imparfaites, ce qui permet de prendre en compte les différentes imperfections des données. De plus, la théorie de l'évidence fournit un cadre formel permettant d'introduire la notion d'ignorance, particulièrement pertinente lorsque les données sont incomplètes. Il s'agit finalement d'une méthode générique puisqu'elle s'appuie sur un ensemble de critères de similarité qui peuvent être définis pour un cas applicatif précis sans remettre en question le cœur de la méthode.

Tout comme pour Samal *et al.* (2004), cette méthode présente l'inconvénient d'évaluer l'ensemble des candidats pour chaque objets géographique d'un jeu de données. Si l'on veut détecter des appariement de cardinalité $N : 1$ ou $N : M$, cette évaluation doit être faite deux fois. Le coût de calcul peut devenir prohibitif lorsque les jeux de données sont importants et denses. De plus, cette méthode ne permet d'apparier que deux jeux de données. Son utilisation nécessiterait donc une adaptation pour gérer un appariement entre plusieurs jeux de données simultanément. Le coût de calcul deviendrait alors rapidement extrêmement important.

Cette approche fondée sur la théorie des croyances propose toutefois une formalisation des connaissances sur les objets géographiques homologues intéressante pour la découverte de relations de filiation entre observations. En particulier, le fait que cette approche envisage l'ignorance parmi les hypothèses de base constitue un aspect intéressant dans le cas d'observations géohistoriques pour lesquelles il est préférable d'envisager des relations de filiation incertaines. Nous proposons donc de reprendre cette approche d'appariement dans le formalisme des fonctions de croyance tout en l'adaptant pour permettre l'appariement de $N (>2)$ sources, la détection de liens d'appariement de cardinalités complexes ($1 : N, N : 1, N : M$) et un meilleur passage à l'échelle.

Nous proposons pour cela de modéliser le problème de la découverte des relations de filiation entre des observations géohistoriques issues d'un ensemble de sources sous la forme d'un problème d'optimisation discrète s'appuyant sur la théorie des croyances. Avant de présenter notre proposition de résolution par optimisation, nous introduisons les éléments les plus importants de la théorie des croyances. Pour de plus amples détails, le lecteur peut se rapporter à (Olteanu, 2008), ou encore (Martin, 2005).

6.2.2 Éléments de la théorie des croyances

Soit $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ un ensemble de propositions appelé **cadre de discernement**. Une **masse de croyance**, ou **fonction de masse**, $m : 2^\Omega \rightarrow [0, 1]$ est une fonction s'exprimant sur l'ensemble $2^\Omega = \{\emptyset, \omega_1, \omega_2, \{\omega_1, \omega_2\}, \dots, \Omega\}$, nommé *power set*, composé de toutes les disjonctions possibles des éléments de Ω . Une fonction de masse exprime un degré de croyance dans une proposition de Ω (appelée proposition singleton) ou dans une disjonction de propositions. Notons que les propositions

doivent être exclusives. Une fonction de masse doit vérifier :

$$\sum_{A \in 2^\Omega} m(A) = 1$$

La représentation de l'incertitude s'effectue alors en attribuant des masses de croyance à des disjonctions d'hypothèses. Un élément $A \in 2^\Omega$ pour lequel $m(A) > 0$ est appelé **élément focal**.

Une *fonction de croyance*, ou *fonction de crédibilité* est une fonction notée $bel : 2^\Omega \rightarrow [0, 1]$ permettant de modéliser la certitude dans la réalité d'une proposition ou d'une disjonction de propositions. Les fonctions de croyance se définissent facilement à partir des masses de croyance. Pour toute proposition $A \in 2^\Omega$, on a :

$$bel(A) = \sum_{B \subseteq A, B \neq \emptyset} m(B)$$

La crédibilité d'une proposition est une mesure relativement pessimiste dans la mesure où seules les connaissances plus précises que A augmentent la croyance en A . D'autres fonctions plus lâches existent. Ainsi, la plausibilité d'une proposition A mesure la force avec laquelle on ne doute pas de A . Elle est définie par :

$$pl(A) = \sum_{B \cap A, B \neq \emptyset} m(B)$$

Enfin, une fonction intermédiaire nommée **probabilité pignistique**, introduite par Smets (1989) effectue un compromis entre crédibilité et plausibilité. Elle consiste à équi-répartir les masses de croyances des disjonctions sur les propositions singleton qui les composent. Cette probabilité s'écrit :

$$P(A) = \sum_{A \subseteq B, B \in 2^\Omega} \frac{m(B)}{|B|(1 - m(\emptyset))}$$

où $m(\emptyset)$ désigne le conflit (voir ci-après).

Plusieurs masses de croyance peuvent être définies sur 2^Ω . Par exemples, plusieurs sources d'informations peuvent se prononcer sur les propositions du cadre de discernement ou leurs disjonctions. Si l'on veut connaître la croyance globale dans une proposition, il faut alors fusionner les différentes masses de croyance. Le fait d'avoir plusieurs sources d'informations peut créer des situations de conflit dans lesquelles les sources proposent des masses de croyances contradictoires. Plusieurs opérateurs de fusion des masses de croyance existent et diffèrent par la stratégie adoptée pour prendre en compte les cas de conflits. Deux stratégies prédominent, nommées respectivement *monde fermé* et *monde ouvert*. La stratégie du monde fermé consiste à considérer que le cadre de discernement est exhaustif. Dans ce cas, le conflit est redistribué sur les éléments de 2^Ω . La stratégie du monde ouvert considère quant à elle un cadre de discernement non exhaustif : la proposition la plus crédible n'est peut être pas représentée. Dans ce cas, le conflit est affecté à un élément de 2^Ω . Ainsi, Smets (1990) propose de placer le conflit sur la proposition vide \emptyset , et Yager (1987) sur Ω .

Opérateurs de fusion

Nous présentons ici les deux opérateurs de fusion les plus utilisés. L'opérateur conjonctif de Dempster (Shafer, 1986) considère un monde fermé. Pour un ensemble de masses de croyances $\{m_1, m_2, \dots, m_n\}$ définies par n sources d'informations sur les éléments de 2^Ω , cet opérateur est défini par :

$$\begin{aligned} m(A) &= \frac{1}{1-k} \sum_{B_1 \cap B_2 \cap \dots \cap B_n = A} \prod_{i=1}^n m_i(B_i) \text{ si } A \neq \emptyset \\ m(\emptyset) &= 0 \end{aligned} \quad (6.1)$$

avec $k = \sum_{B_1 \cap B_2 \cap \dots \cap B_n = \emptyset} \prod_{i=1}^n m_i(B_i)$.

$m(A)$ désigne la masse de croyance fusionnée de m_1, \dots, m_n . La valeur k est la **mesure de conflit** entre les diverses sources d'informations. L'opérateur de Dempster effectue une normalisation des croyances par le conflit. Cette normalisation a pour effet d'augmenter artificiellement les croyances dans les propositions pour lesquelles les sources d'informations sont concordantes.

Un second opérateur, proposé par Smets (1990) reprend celui de Dempster dans le cas d'un monde ouvert. La différence avec l'opérateur conjonctif est qu'au lieu de normaliser les croyances, le conflit est déplacé sur $m(\emptyset)$. On a alors :

$$\begin{aligned} m(A) &= \sum_{B_1 \cap B_2 \cap \dots \cap B_n = A} \prod_{i=1}^n m_i(B_i) \text{ si } A \neq \emptyset \\ m(\emptyset) &= k \end{aligned} \quad (6.2)$$

Cet opérateur suppose que les sources d'informations sont totalement fiables. Lorsque ce n'est pas le cas, la stratégie généralement choisie est de redistribuer le conflit sur les propositions impliquées dans l'émergence de ce conflit, c'est-à-dire dont les sources d'informations sont contradictoires (voir (Lefevre *et al.*, 2002)).

Affaiblissement

La théorie des croyances permet également de modéliser la fiabilité des sources d'information lorsque cette fiabilité est quantifiable. Le manque de fiabilité d'une source résulte en un *affaiblissement* de la croyance qu'elle place en une hypothèse et un accroissement de l'ignorance¹². L'affaiblissement α_i de la source d'information i s'exprime en écrivant sa fonction de masse de la façon suivante :

$$\begin{aligned} m_i(A) &= \alpha_i m_i(A), A \in 2^\Omega \\ m_i(\Omega) &= 1 - \alpha_i(1 - m_i(A)) \end{aligned} \quad (6.3)$$

Décision

La fusion des masses de croyance permet d'obtenir des masses de croyances globales résumant l'avis des différentes sources d'informations sur les différentes propositions. Il reste alors à prendre une décision, c'est-à-dire à choisir une proposition (généralement un singleton). Différents critères de décision existent, les plus classiques consistent à sélectionner la proposition de crédibilité, de plausibilité ou de probabilité pignistique maximale.

12. La somme des croyances de cette source est donc toujours égale à 1.

6.2.3 Approche par optimisation discrète

Les bases de la théorie des croyances ayant été exposées, nous introduisons maintenant notre approche d'appariement multicritère fondée sur cette théorie. Pour ce faire, nous formulons ce problème d'appariement comme problème d'optimisation discrète.

Soit \mathcal{H} l'ensemble formé de tous les hypergraphes admissibles pouvant être construits à partir de tous les graphes de filiation admissibles possibles liant les observations de O .

Nous introduisons une fonction de **vraisemblance d'un hypergraphe de filiation** $V : \mathcal{H} \rightarrow \mathbb{R}^+$ qui synthétise toutes les vraisemblances des arcs d'un hypergraphe $H_f \in \mathcal{H}$. Cette fonction mesure à quel point un hypergraphe admissible est une description vraisemblable des transformations ayant eu lieu entre les entités du monde réel représentées par les observations de O .

Le problème revient à rechercher l'hypergraphe H_f^* de vraisemblance maximale au sein de tous les hypergraphes admissibles exprimables sur un ensemble d'observations. Il s'agit donc de déterminer :

$$H_f^* = \operatorname{argmax}\{V(H_f) : H_f \in \mathcal{H}\} \quad (6.4)$$

Dans le reste de cette section, nous définissons précisément cette fonction de vraisemblance d'un hypergraphe de filiation. Celle-ci est intrinsèquement liée à la fonction de vraisemblance d'un hyperarc Φ que nous avons définie pour l'instant de façon générique. Nous procédons par étapes, en précisant dans un premier temps Φ puis en remontant jusqu'à l'expression de V afin d'en donner l'expression détaillée.

Passage à la théorie des croyances

Soit $H_f(O \cup o_\emptyset, E, S, \mathcal{T})$ un hypergraphe de filiation admissible dont on cherche à savoir s'il représente un ensemble de processus spatio-temporels vraisemblables. Nous définissons deux hypothèses sur chaque arc $e \in E$:

ω_e : e est l'arc décrivant le processus spatio-temporel le plus vraisemblable entre les observations de sa queue et de sa tête,

$\neg\omega_e$: l'hypothèse contraire.

Pour l'ensemble $\{e_1, e_2, \dots, e_n\}$ de tous les arcs de H_f , nous pouvons créer le cadre de discernement Ω contenant toutes les hypothèses sur ces arcs :

$$\Omega = \{\omega_{e_1}, \neg\omega_{e_1}, \dots, \omega_{e_n}, \neg\omega_{e_n}\}$$

Les sources d'informations qui vont croire plus ou moins à ces hypothèses sont, pour nous, des critères portant sur l'identité et la succession temporelle des observations de la queue et de la tête de chaque hyperarc. Nous considérons que la croyance en un hyperarc ne dépend pas de celle des autres arcs de H_f . Autrement dit, nous traitons chaque arc comme un sous-problème du problème général de détermination des processus spatio-temporels les plus vraisemblables. Pour ce faire, nous utilisons la notion de *source spécialisée*, introduite par Appriou (1991) et utilisée par (Olteanu, 2008). L'idée est de décomposer les sources d'informations en sources se spécialisant

sur chaque hypothèse singleton dans 2^Ω . Plus exactement, chaque source spécialisée n'a que trois éléments focaux : une hypothèse singleton $\omega \in \Omega$, son contraire $\neg\omega$ et la disjonction totale Ω . Ainsi, une source d'information peut croire qu'une hypothèse singleton est vraie, qu'elle est fausse ou bien peut être dans l'incertitude totale. Sur chaque source spécialisée est également définie un coefficient de fiabilité α permettant d'affaiblir les sources peu fiables.

Pour nous, les sources d'informations vont se spécialiser sur chaque hyperarc. Étant donné un nombre M de sources d'informations et n hyperarcs, cela implique de définir $M.n$ fonctions de masse de croyance de la forme $m_{ie_j}(A)$, $A \in \{\omega_{e_j}, \neg\omega_{e_j}, \Omega\}$. Dans son travail, Appriou (1991) s'intéresse à de l'apprentissage statistique et distribue les masses de croyances sur ω et $\neg\omega$, la croyance incertaine (Ω) étant due au manque de fiabilité des sources d'informations. Il introduit les deux modèles de fonction de masses suivants :

Modèle 1 :

$$\begin{cases} m_{ij}(\omega) = \alpha_{ij}M_{ij} \\ m_{ij}(\neg\omega) = \alpha_{ij}(1 - M_{ij}) \\ m_{ij}(\Omega) = 1 - \alpha_{ij} \end{cases}$$

Modèle 2 :

$$\begin{cases} m_{ij}(\omega) = 0 \\ m_{ij}(\neg\omega) = \alpha_{ij}(1 - M_{ij}) \\ m_{ij}(\Omega) = 1 - \alpha_{ij}(1 - M_{ij}) \end{cases}$$

où M_{ij} est la vraisemblance statistique de la proposition ω et α_{ij} un coefficient d'affaiblissement de la source focalisée ij . Si les sources d'informations ne sont pas fiables, toute la croyance est donc placée dans l'ignorance.

Dans notre cas, la valeur de M_{ij} est la valeur des mesures d'identité ou de succession temporelle. De plus, nous considérons que toutes les sources focalisées du même type (antécédence, sauts, etc.) ont la même fiabilité. Enfin, les deux modèles peuvent être interprétés comme des visions plus ou moins prudentes du problème. Le premier modèle est peu prudent, les sources d'informations s'exprimant sur la validité ou la non-validité d'une proposition : si la source ne croit pas en ω , elle considère cette proposition fausse. Le second modèle est nettement plus prudent : tout ce qui n'appuie pas la proposition ω est placé dans l'ignorance. Ainsi, si la source ne croit pas en ω , elle est dans une situation d'incertitude.

Un tel comportement est particulièrement intéressant pour notre cas d'application car il permet de traiter différemment les informations sur l'identité et sur la succession temporelle. Les mesures d'identité doivent s'exprimer sur l'existence d'une filiation entre deux observations. Pour une mesure donnée, l'existence d'une filiation est crédible ou non : elles suivent le premier modèle. Les sources d'informations sur la succession temporelle suivent quant à elles le second modèle. En effet, les mesures de succession permettent de savoir si une filiation est envisageable car logique temporellement, mais elles ne permettent pas de savoir si cette filiation est pertinente. Finalement, les deux modèles de Appriou (1991) peuvent être exprimés pour la découverte des relations de filiation de la façon suivante :

Pour les mesures d'identité :

$$\begin{cases} m_{ie_j}(\omega_e) = \alpha_i c \\ m_{ie_j}(\neg\omega_e) = \alpha_i(1 - c) \\ m_{ie_j}(\Omega) = 1 - \alpha_i \end{cases}$$

Pour les mesures de succession temporelle :

$$\begin{cases} m_{ie_j}(\omega_e) = 0 \\ m_{ie_j}(\neg\omega_e) = \alpha_i(1 - c) \\ m_{ie_j}(\Omega) = 1 - \alpha_i(1 - c) \end{cases}$$

où c est la valeur de la mesure calculée pour l'hyperarc e_j .

Fusion des croyances sur un hyperarc

L'ensemble des mesures de similarité d'identité et de succession temporelle s'expriment sur chaque hyperarc de H_f sous la forme de fonctions de masses. Chaque mesure donne une valeur de croyance et il faut donc fusionner l'ensemble des masses de croyances sur les hypothèses définies pour l'arc pour obtenir la fonction de masse totale m_{e_j} pour chacune des hypothèses. Pour cela il faut choisir un opérateur de fusion et donc décider si l'on se place dans le cas d'un *monde ouvert* ou *fermé*.

Évaluer la vraisemblance d'un hyperarc revient à évaluer la vraisemblance du processus spatio-temporel sous-jacent liant ses observations. Cette évaluation se fait sur les trois hypothèses définies précédemment, or d'autres relations de filiations entre les observations membres de l'hyperarc pourraient tout à fait conduire à un autre hyperarc plus vraisemblable que celui en cours d'évaluation. Nous nous trouvons donc dans le cas d'un monde ouvert, où d'autres hypothèses que celles du cadre de discernement peuvent exister.

Nous utilisons donc l'opérateur conjonctif non normalisé proposé par Smets (1990) :

$$\begin{cases} m_{e_j}(A) = \sum_{B_1 \cap \dots \cap B_m = A \neq \emptyset} \prod_{i=1}^M m_{ie_j}(B_i), \\ m_{e_j}(\emptyset) = \sum_{B_1 \cap \dots \cap B_m = \emptyset} \prod_{i=1}^M m_{ie_j}(B_i), \end{cases}$$

$m_{e_j}(\emptyset)$ représente le conflit émergent des sources d'information contradictoires. Dans notre cas, le conflit apparaît lorsque les identités des observations agrégées de la queue et de la tête de l'hyperarc évalué sont similaires mais que la relation de filiation est soit dans le sens rétrotemporel, soit qu'il existe d'autres sources géohistoriques intercalées entre celles contenant la queue et la tête de l'arc. Autrement dit, cela signifie que les sources optimistes croient fortement dans la vraisemblance de l'arc tandis que les sources pessimistes croient dans l'hypothèse contraire.

Une fois les masses de croyance fusionnées pour chaque arc $e \in E$, nous souhaitons évaluer la vraisemblance de celui-ci. Pour cela, nous pouvons évaluer la crédibilité, la plausibilité ou la probabilité pignistique de l'hypothèse ω_e . La probabilité pignistique est dans notre cas particulièrement pertinente. En effet, c'est une mesure de compromis bien adaptée au cas où la crédibilité et la plausibilité sont très différentes, ce qui est particulièrement le cas lorsque l'ignorance est élevée. La crédibilité

est alors faible tandis que la plausibilité est proche de 1. Or, les mesures optimistes et pessimistes peuvent aboutir à une forte ignorance. Pour tout hyperarc $e \in E$, la probabilité pignistique de l'hypothèse ω_e "l'arc est vraisemblable" est :

$$P(\omega_e) = \frac{m_e(\omega_e)}{1 - m(\emptyset)} + \frac{m(\Omega)}{2(1 - m(\emptyset))}$$

On a finalement, pour tout hyperarc $e \in E$:

$$\Phi(e) = \phi(o_Q, o_T) = P(\omega_e) \quad (6.5)$$

Vraisemblance des hyperarcs décrivant des apparitions et disparitions

Les hyperarcs correspondant à des processus spatio-temporels d'apparition et de disparition sont de la forme $e = (o_\emptyset, o)$ ou $e = (o, o_\emptyset)$ et connectent l'observation o à l'observation nulle o_\emptyset . Nous venons de définir la vraisemblance de tels hyperarcs comme une mesure fusionnée à partir de critères comparant deux observations, or la comparaison avec l'observation nulle n'a pas de sens. La théorie des croyances offre en réalité un cadre permettant aisément de modéliser le cas de ces hyperarcs. En effet, lorsqu'il y a apparition ou disparition, aucune observation ne correspond à o : toutes les sources d'informations sont alors dans une situation d'ignorance totale. La totalité de la masse de croyance est alors placée sur Ω .

La vraisemblance Φ d'hyperarcs représentant des apparitions ou disparitions est alors :

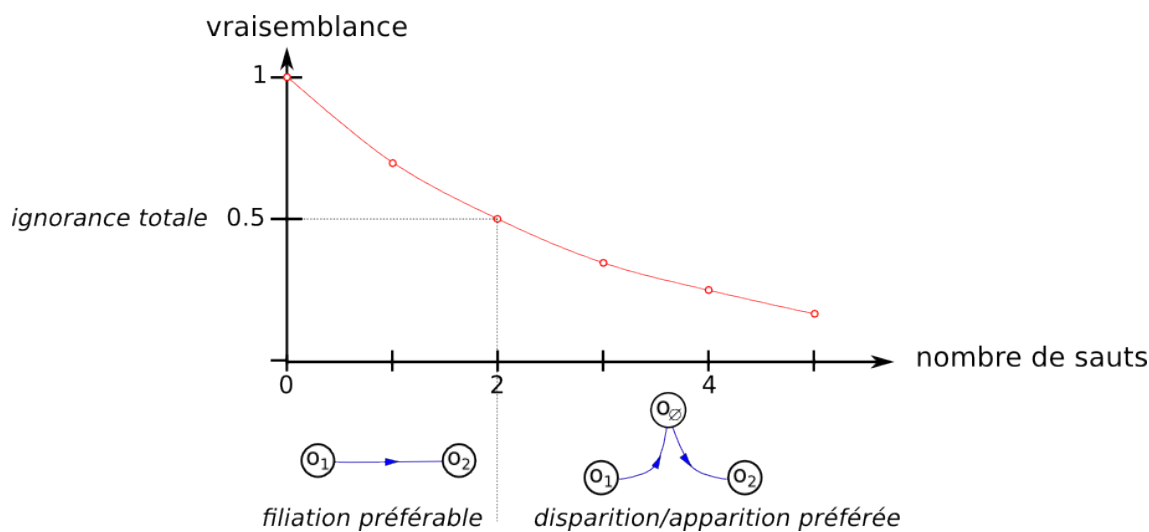
$$\begin{aligned} \forall o \in O, \forall e = (o, o_\emptyset) \text{ ou } e = (o_\emptyset, o), \\ \Phi(e) = \frac{1}{2} \end{aligned} \quad (6.6)$$

En effet, dans ce cas, $m_e(\Omega) = 1$ et le conflit est nul. La probabilité pignistique de ω_e vaut alors $\frac{1}{2}$.

Paramétrage des mesures

La probabilité d'une apparition ou d'une disparition est donc indépendante des mesures ou de la nature des observations, ce qui permet un paramétrage simple des mesures de similarité d'identité et de succession temporelle. En effet, l'existence d'un processus spatio-temporel de cardinalité supérieure à 1 : 0 ou 0 : 1 est vraie, pour une mesure donnée, si celle-ci est au moins supérieure à $\frac{1}{2}$. En dessous, la mesure considère que les observations ne sont pas liées.

Ainsi, nous avons présenté (voir le paragraphe 6.1.2) la mesure de succession fondée sur le rang des sources, qui dépendait d'un paramètre ϵ . Nous avons dit que la détermination de ce paramètre dépendait d'une valeur limite entre vraisemblance et non vraisemblance d'une relation de filiation. Cette valeur correspond à la vraisemblance d'une apparition ou d'une disparition. En effet, si une relation est jugée peu vraisemblable, alors il est préférable de ne pas lier les observations. Nous n'avons introduit pour l'instant un paramètre uniquement pour la mesure c_{rank} , mais toutes les mesures doivent être paramétrées de la même façon. En particulier, les mesures de similarité d'identités doivent être paramétrables afin qu'elles puissent s'adapter aux spécificités des observations qu'elles évaluent. La mesure c_{ant} valant déjà $\frac{1}{2}$ dans les cas d'incertitude sur l'antécédence (voir le chapitre 4), elle ne nécessite pas de

FIGURE 6.5 – Exemple de paramétrisation de la mesure de succession temporelle c_{rank} .

paramétrage supplémentaire.

La façon habituelle de déterminer une limite entre un appariement acceptable ou non pour un critère donné est de fixer un seuil au delà duquel l'appariement de deux objets est rejeté. Par exemple, si l'on cherche les homologues de deux bases de données composées d'objets ponctuels à l'aide d'une distance euclidienne, le seuil de cette distance pourra être la somme des précisions planimétriques de chaque base de données. Au delà de cette limite les objets sont arbitrairement jugés non homologues.

La détermination de ϵ , pour chaque mesure, dépend de la nature de celle-ci. Cependant, il est important que cette détermination soit simple pour qu'un utilisateur construisant un graphe de filiations soit capable de maîtriser le comportement des mesures qu'il utilise. En particulier, le sens de la valeur de ϵ fixée par l'utilisateur doit être intuitif.

Pour c_{rank} , le paramètre ϵ peut être écrit $\epsilon = \frac{-s_{max}}{\log(0.5)}$. s_{max} est alors le **nombre de sauts maximum** avant de considérer une relation comme non vraisemblable. La figure 6.5 illustre le comportement de c_{rank} lorsque l'on considère qu'une relation de filiation devient non vraisemblable si elle "saute" au dessus de deux sources géohistoriques.

Vraisemblance d'un hypergraphe de filiation

Maintenant que nous avons défini Φ , nous pouvons construire la fonction de vraisemblance V d'un hypergraphe. Pour cela, nous considérons que les processus décrits par les hyperarcs sont indépendants les uns des autres¹³. Cela signifie que l'agrégation de deux processus spatio-temporels ne conduit pas à la création d'un nouveau processus plus vraisemblable étant donné les connaissances dont on dispose sur les observations et les transformations de l'espace. Le produit des probabilités pignistiques de tous les arcs d'un hypergraphe fournit donc la probabilité que cet

13. Ceci n'est vrai que parce que l'on ne traite de processus spatio-temporels que l'on a choisi de représenter comme élémentaires.

hypergraphe soit vraisemblable. Le nombre d'arcs d'un hypergraphe peut être élevé, ce qui pose des problèmes de calculs de ce produit. Une façon habituelle de contourner ce problème consiste à calculer la somme des logarithmes de ces probabilités. Le fait de rechercher l'hypergraphe le plus vraisemblable par optimisation signifie qu'il faut être capable de comparer la vraisemblance de deux hypergraphes. Or, le nombre arcs variant d'un hypergraphe à l'autre, ils ne sont pas comparables. Une façon simple de permettre cette comparaison est d'attribuer la probabilité de chaque arc aux observations qui le composent. Ainsi, la vraisemblance d'un hypergraphe devient la probabilité pour que chacune de ses observations appartienne au processus spatio-temporel décrivant le plus vraisemblablement la transformation de cette observation.

Nous pouvons finalement exprimer $V(H_f)$ pour tout hypergraphe $H_f = (O, E, S, \mathcal{T})$ admissible :

$$\log V(H_f) = \sum_{e \in E} \log(\Phi(e)) |e| \quad (6.7)$$

$|e|$ désignant le nombre d'observations simples et non nulles composant e .

Vers un problème de minimisation

Dans la section suivante, nous proposons une méthode de résolution du problème d'optimisation discrète que nous venons de présenter. Pour que cette résolution soit possible, il est nécessaire de transformer le problème de maximisation en un problème de minimisation. Le passage peut être effectué en mesurant non plus la vraisemblance d'une hyperarc ω_e mais l'hypothèse inverse. Il s'agit finalement de minimiser l'invraisemblance d'un hypergraphe. Ce passage peut être réalisé en modifiant seulement la fonction V qui devient :

$$\log Iv(H_f) = \sum_{e \in E} \log(1 - \Phi(e)) |e| \quad (6.8)$$

Le problème d'optimisation devient alors :

$$\operatorname{argmin}\{\log Iv(H_f) : H_f \in \mathcal{H}\} \quad (6.9)$$

Formalisé ainsi, un hypergraphe de filiation peut être construit si l'on dispose des éléments suivants :

- un ensemble d'observations O dont les schémas sont alignés,
- une définition de l'identité des observations de O ,
- un ensemble de mesures évaluant la similarité des attributs composant les identités des observations,
- un coefficient de fiabilité associé à chaque mesure ;

6.2.4 Un recuit simulé mono-objectif pour découvrir H_f^*

Résoudre le problème de découverte des relations de filiation entre observations revient à trouver le minimum global de la fonction objectif présentée dans le paragraphe précédent. À ce minimum global correspond le graphe H_f^* décrivant les processus spatio-temporels les plus vraisemblables entre les différentes observations. Afin de déterminer ce graphe il faut explorer un espace de recherche constitué de

solutions possibles et à choisir la solution minimisant la fonction objectif. L'énumération complète des solutions possibles n'étant pas envisageable, l'approche usuelle consiste alors à explorer cet espace de recherche de façon efficace. Parmi les différentes méthodes existantes, la classe des **métaheuristiques** permet de résoudre des problèmes d'optimisation pour lesquels *on ne connaît pas d'algorithmes permettant de repérer un optimum global à coup sûr et en un nombre fini de calculs* (Dréo *et al.*, 2003). Nous sommes typiquement dans un tel cas, en particulier lorsque plusieurs ordres temporels sont possibles entre observations.

Méthodes de résolution de problèmes d'optimisation combinatoire

Résoudre un problème d'optimisation discrète pour lequel on connaît une fonction objectif consiste à trouver, au sein d'un ensemble discret et fini de solutions appelé **espace de recherche**, la **solution optimale** minimisant (ou maximisant, selon le problème) cette fonction objectif. Lorsque l'espace de recherche est très grand, la difficulté est alors de trouver la solution optimale du problème dans un temps raisonnable. L'idée générale de toutes les méthodes de résolution de problèmes d'optimisation est de parcourir l'espace de recherche du problème de façon efficiente ; seule la stratégie adoptée diffère. Deux grandes classes de méthodes existent :

1. Les méthodes dites exactes qui consistent à parcourir entièrement l'espace de recherche. Afin d'accélérer le parcours, ces méthodes cherchent à détecter au plus tôt les zones de l'espace de recherche ne contenant pas la solution optimale pour les éliminer. La programmation dynamique et la méthode *branch and bound* sont les deux représentantes les plus connues de cette classe de méthodes. L'avantage majeur des méthodes exactes est qu'elles assurent de trouver la solution optimale du problème.
2. Les méthodes approchées sont utilisées lorsque le problème est de grande taille et que l'on peut se contenter d'une solution proche de l'optimum. Le principe des méthodes approchées est d'éviter un parcours exhaustif de l'espace de recherche à l'aide d'heuristiques diverses. Parmi les méthodes approchées, la classe des **métaheuristiques** rassemble des méthodes génériques qui offrent une stratégie générale permettant de guider la recherche d'une solution optimale pour un problème d'optimisation spécifique. Plutôt de méthodes "clé en main", il s'agit de cadres de conception de méthodes approchées pour un problème donné. Le schéma générale d'une métaheuristique consiste à choisir¹⁴ une ou plusieurs solutions dans l'espace de recherche, puis à parcourir cet espace en appliquant de légères modifications aux solutions initiales pour obtenir de nouvelles solutions, de nouveau modifiées, jusqu'à atteindre une solution jugée suffisamment proche de l'optimum global. Les métaheuristiques se répartissent selon le nombre de solutions qu'elles traitent. Ainsi, les métaheuristiques **globales** (ou distribuées) tentent de parcourir l'espace de recherche avec plusieurs solutions en parallèle, tandis que les méthodes **locales** (ou de voisinage) ne s'appuient que sur une solution. S'il existe diverses métaheuristiques (voir (Dréo *et al.*, 2003)), le **recuit simulé** (local) et les **algorithmes évolutionnaires** (globaux) sont les plus généralement utilisés. Le principale

14. Généralement de manière aléatoire

avantage des méthodes approchées est leur capacité à traiter de problèmes de très grande taille. Quant aux métaheuristiques, elles fournissent un cadre de résolution générique qui leur permet d'être appliquées à des problèmes variés ou difficiles à définir formellement. La non-optimalité de ces méthodes est cependant leur défaut majeur. De plus, elles demandent de définir des heuristiques - notamment pour le parcours de l'espace de recherche- qui peuvent être difficiles à maîtriser.

Choix d'une méthode

Le premier choix à effectuer porte sur l'utilisation d'une méthode exacte ou approchée. Si l'utilisation de méthodes exactes assurerait de trouver l'optimum global, elles sont mal adaptées à notre problème. Tout d'abord, la taille de l'espace de recherche devient importante lorsque le nombre de sources et d'observations géohistorique augmente et nous ne disposons pas de critères permettant d'éviter le parcours de la majeure partie de l'espace de recherche. De plus, notre objectif est de mettre en place une méthode générique pouvant être adaptée à des observations de différentes natures. Bien que toutes les mesures (d'identité et temporelles) soient synthétisées sous la forme de croyance, la forme des graphes de filiation peut varier selon la nature des observations¹⁵. Par conséquent, la forme de l'espace de recherche varie également selon les observations considérées, ce qui gêne d'autant plus la mise en place de critères évitant un parcours exhaustif. Enfin, l'optimalité n'est pas une nécessité absolue, d'autant plus que le résultat nécessitera le plus souvent des corrections manuelles, les observations et les mesures ne transcrivant pas toute la complexité des phénomènes représentés. Nous nous dirigeons donc vers une méthode approchée.

Parmi les méthodes approchées, le cadre générique offert par les métaheuristiques est particulièrement pertinent pour les raisons que nous venons d'énoncer. De nombreux travaux ont porté sur la comparaison entre métaheuristiques globales et locales pour différentes classes de problèmes (Manikas et Cain, 1996; Rossi-Doria *et al.*, 2003; Fouskakis et Draper, 2002; Kohonen, 1999) et ont montré qu'il n'y a pas de règle claire permettant de connaître a priori la méthode la mieux adaptée à un problème donné, chacune ayant des avantages et des manques dont les effets sont difficiles à estimer. Fouskakis et Draper (2002) ont toutefois montré que l'hybridation de deux métaheuristiques était une piste intéressante qui permet de combler les manques des méthodes ainsi combinées. Les principales méthodes hybrides peuvent être trouvées dans (Talbi, 2002).

Dans son article, Kohonen (1999) rapproche les deux métaheuristiques les plus utilisées, le recuit simulé et l'algorithme génétique (membre des méthodes évolutionnaires), et montre que leurs apparentes différences cachent en réalité des comportements relativement similaires. En effet, tous deux utilisent un parcours de l'espace de recherche par voisinage, supposant ainsi que de meilleures solutions existent à proximité des solutions connues. Il montre par ailleurs que lorsque cette hypothèse se vérifie, une optimisation locale est avantageuse. La méthode du recuit simulé est donc préférable quand le problème peut être vu comme une suite d'optimisations locales.

15. Par exemple, des bâtiments n'évoluent pas comme des parcelles. Les premiers auront tendance à apparaître et disparaître tandis que que les secondes se réorganisent.

La découverte des relations de filiation est un problème d'optimisation intrinsèquement local. En effet, chaque processus spatio-temporel ne concerne qu'un petit nombre d'observations et n'occupe qu'une petite partie du graphe total. **Nous choisissons donc le recuit simulé comme méthode de résolution du problème de découverte des relations de filiations d'un graphe géohistorique.**

Principe du recuit simulé

Le recuit simulé est une métaheuristique locale dont le principe consiste à se déplacer itérativement de voisins en voisins dans l'espace de recherche en sélectionnant à chaque étape un voisin de la solution courante comme nouvelle solution selon une distribution de probabilité qui dépend de la qualité des différents voisins (Kirkpatrick *et al.*, 1983). Le schéma du recuit simulé est illustré en figure 6.6 pour le cas de la découverte des relations de filiation. Ainsi, les meilleurs voisins¹⁶ ont une plus grande probabilité d'être choisis comme nouvelle solution courante. Afin d'éviter d'être piégé dans des minima locaux de la fonction objectif, le recuit simulé comporte un paramètre T nommé **température** dont le rôle est le suivant :

- si T est grand, tous les voisins d'une solution ont environ la même probabilité d'être sélectionnés comme nouvelle solution,
- si T est petit, la probabilité de choisir un voisin de mauvaise qualité est faible,
- si T est nul, seuls les voisins améliorant la fonction objectif ont une probabilité non nulle d'être choisis.

La température d'un recuit simulé décroît avec le temps selon une loi empiriquement déterminée. Plusieurs lois de décroissance de la température ont été proposées, les principales pouvant être trouvées dans (Nourani et Andresen, 1998). Les lois de décroissance linéaires, géométriques et quadratiques sont les plus couramment utilisées.

Si l'on note E la fonction objectif (aussi appelée énergie), x la solution courante, x' une solution voisine et $\Delta E = E(x') - E(x)$ la différence d'énergie entre les deux solutions, la probabilité d'accepter x' comme nouvelle solution courante est généralement donnée par le **critère de Metropolis** :

$$p(\Delta E) = e^{(-\frac{\Delta E}{T})} \quad (6.10)$$

La solution x' est alors acceptée avec une certaine probabilité pouvant être non nulle même si elle dégrade la fonction objectif si la température est élevée. Le recuit simulé poursuit sa recherche jusqu'à atteindre un critère d'arrêt qui peut être :

- un nombre maximal d'itérations,
- un critère de stagnation (temps ou nombre d'itérations sans amélioration significative de la solution),
- un temps maximal.

Le recuit simulé est une méthode simple, dont la convergence théorique vers la solution optimale a été démontrée et étudiée dans de nombreux travaux (voir (Henderson *et al.*, 2003)). Cependant, plusieurs de ses paramètres doivent être fixés de manière empirique :

- la loi de décroissance de température,

16. Le recuit simulé étant classiquement adapté à des problèmes de minimisation, un voisin est meilleur si la valeur de la fonction objectif est plus faible.

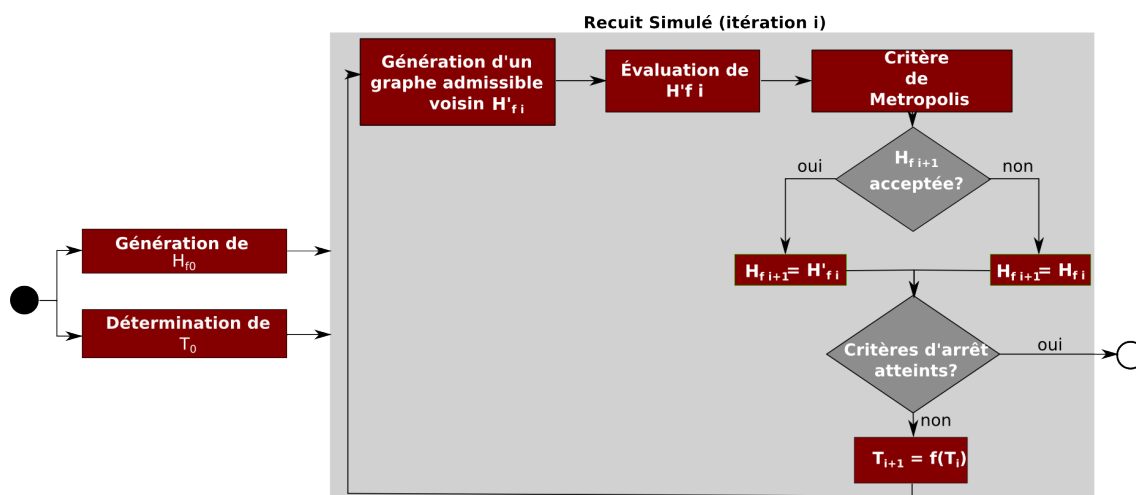


FIGURE 6.6 – Schéma de l'algorithme de recuit simulé pour la découverte des relations de filiation.

- la température initiale d'une optimisation,
- les critères d'arrêt.

La détermination de ces paramètres est une étape cruciale car elle conditionne la convergence du recuit simulé. En effet, si la température reste trop élevée durant toute l'optimisation, le recuit simulé sera équivalent à une marche aléatoire. À l'inverse si sa décroissance est trop rapide, le risque est alors de se bloquer dans un minimum local. Une température initiale trop haute ou trop basse générera les mêmes effets. Quant aux critères d'arrêts, ils doivent être déterminés de façon à aboutir à un compromis entre temps de calcul et convergence de l'algorithme.

Application à l'identification des filiations

Nous avons choisi d'utiliser la méthode du recuit simulé pour réaliser la phase de découverte des relations de filiations au sein d'un ensemble d'observations. Le schéma 6.6 présente le déroulement général d'un algorithme de recuit simulé adapté à notre problème. Dans les paragraphes suivants, nous en détaillons les étapes principales qui doivent être créés pour mettre en place un tel algorithme. En particulier, nous nous concentrons sur les quatre éléments majeurs de la méthode :

- la génération du voisinage d'une solution,
- la détermination de la température initiale,
- le choix d'une loi de décroissance de température,
- la détermination de la solution initiale.

Génération du voisinage d'un hypergraphe de filiations

La stratégie de génération du voisinage d'une solution joue un rôle important dans l'efficacité d'une optimisation par recuit simulé. Pour Alizamir *et al.* (2008), cette stratégie doit être efficace et efficiente. L'efficacité renvoie ici à sa capacité à générer des solutions couvrant l'intégralité de l'espace de recherche. L'efficience désigne sa capacité à se déplacer *de façon adéquate* dans l'espace de recherche, c'est à dire en sortant rapidement des minima locaux tout en atteignant rapidement les

bonnes solutions. Cette efficacité dépend du nombre d'étapes nécessaires pour se déplacer d'une solution à une autre (vitesse), de la taille du voisinage d'une solution, de la quantité d'information modifiée par un déplacement dans l'espace de recherche et de la difficulté de calcul des voisins. S'il est toujours préférable de minimiser le nombre de déplacements nécessaires pour aller d'une solution à une autre, les autres éléments de l'efficacité sont spécifiques au problème, sont parfois contradictoires et nécessitent des compromis (Nourani et Andresen, 1998).

Afin de former une stratégie respectant le mieux possible ces différents aspects, nous générons le voisinage d'un hypergraphe de filiation H_f en appliquant une modification élémentaire à l'un des processus spatio-temporel qu'il décrit afin de créer un nouvel hypergraphe admissible H'_f . Ceci revient en réalité à ajouter ou supprimer une relation de filiation du graphe G_f correspondant à H_f sous les contraintes d'admissibilité décrites dans le paragraphe 6.1.2. Pour transcrire ces deux actions au niveau d'un hypergraphe de filiation, les modifications doivent être effectuées sur les hyperarcs du graphe. Quatre modifications permettent de transcrire l'action d'ajout ou de suppression d'une relation de filiation :

1. **l'extension** d'un hyperarc consiste à étendre un hyperarc en ajoutant une nouvelle observation dans sa queue ou sa tête.
2. **la réduction** d'un hyperarc consiste à retirer une observation de sa queue ou de sa tête.
3. **la fusion** de deux hyperarcs aboutit à la création d'un nouvel hyperarc de cardinalité $n : m$.
4. **la scission** d'un hyperarc aboutit à la création de deux nouveaux hyperarcs.

Puisque nous considérons depuis le début que toute observation o d'un hypergraphe de filiation est toujours connectée à une autre observation ou à l'observation nulle, elle appartient nécessairement à deux hyperarcs $e_1 = (Q_1, T_1)$ et $e_2 = (Q_2, T_2)$ où $o \in T_1$ et $o \in Q_2$. Dès lors, les quatre modifications fixées peuvent s'écrire comme la fusion de deux hyperarcs voisins. Nous définissons le voisinage de deux hyperarcs de la façon suivante : **deux hyperarcs sont voisins si les observations agrégées de leur queues (et resp. de leurs têtes) sont connectées spatialement et sont issues de la même source**. Ainsi, cela signifie que si deux hyperarcs sont voisins, toutes les observations de leurs têtes respectives sont issues de la même source et qu'au moins deux d'entre elles sont connectées spatialement. Il en va de même pour leurs queues. Nous considérons de plus que l'observation nulle est connectée spatialement à toutes les autres, et qu'elle appartient à toutes les sources indistinctement. La figure 6.7 fournit deux exemples de voisinages, l'un pour deux arcs de cardinalité $1 : 1$ et l'un avec un arc comprenant l'observation nulle. Les relations spatiales sont figurées par des liens en vert. Nous fixons alors comme contrainte sur les modifications le fait qu'elles s'appliquent uniquement sur des hyperarcs voisins. Ainsi, un arc ne peut être étendu qu'avec des observations membres d'un arc voisin de celui-ci. Il en va de même pour la modification de fusion. Nous plaçons cette contrainte de façon à éviter de créer des hyperarcs ne traduisant aucun processus spatio-temporel crédible. Par exemple, il est illogique que deux parcelles éloignées de plusieurs kilomètres se trouvent fusionnées en une seule parcelle.

Exprimées de cette façon, les quatre modifications peuvent toutes être décrites

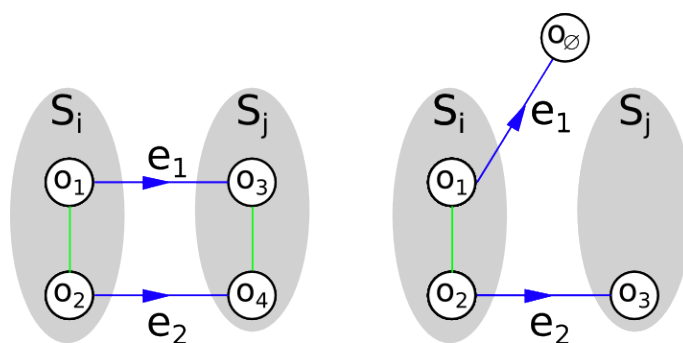


FIGURE 6.7 – Deux cas d’hyperarcs voisins

comme des opérations d’union d’hyperarcs voisins ou de division d’un hyperarc en deux arcs voisins. La figure 6.8 illustre ainsi les opérations équivalentes à chaque modification sur des exemples.

Création d’hypergraphes admissibles

L’objectif étant de créer des hypergraphes admissibles, les modifications doivent être effectuée de façon à créer de nouveaux hyperarcs respectant les contraintes C_1 , C_2 et C_3 fixées dans le paragraphe 6.1.2. Ceci est surtout valable pour les modifications d’extension et de fusion qui créent ajoutent des observation à des arcs existants. Il est possible de vérifier si une telle modification aboutit à un hypergraphe admissible en éliminant lors de la création des modifications d’extension et de fusion les hyperarcs voisins dont l’agrégation aboutirait à un nouvel arc violant les contraintes d’admissibilité. Il y a donc préalablement à la création de toute modification un filtrage des hyperarcs candidats, de façon à assurer la production de nouveaux hypergraphes admissibles. Ce filtrage consiste à vérifier qu’une fois la modification appliquée à un hyperarc e , l’hyperarc résultat e' respecte les contraintes C_1 , C_2 et C_3 définies sur tout hypergraphe de filiation. La procédure générale de génération d’une modification d’un hyperarc quelconque e est décrite dans le schéma 6.9 et se compose des étapes suivantes :

1. Une étape de filtrage des hyperarcs candidats afin de créer un nouvel hypergraphe admissible.
2. Après filtrage, un candidat est choisi aléatoirement. S’il ne reste aucun candidat, la création de la modification échoue.
3. Enfin, la modification est créée. Une modification peut être stockée en conservant l’hyperarc modifié qu’il faudra supprimer, et en créant les nouveaux arcs résultant de la modification.

Les quatre modifications possibles et les contraintes qui s’appliquent sur elles sont détaillés dans les paragraphes suivants.

Extension d’un hyperarc

Étendre un hyperarc e consiste fusionner un hyperarc avec un arc voisin de cardinalité $0 : 1$ ou $1 : 0$. Deux cas de figure peuvent se présenter : soit l’hyperarc est un appariement de cardinalité supérieure ou égale à $1 : 1$, soit il décrit un apparition

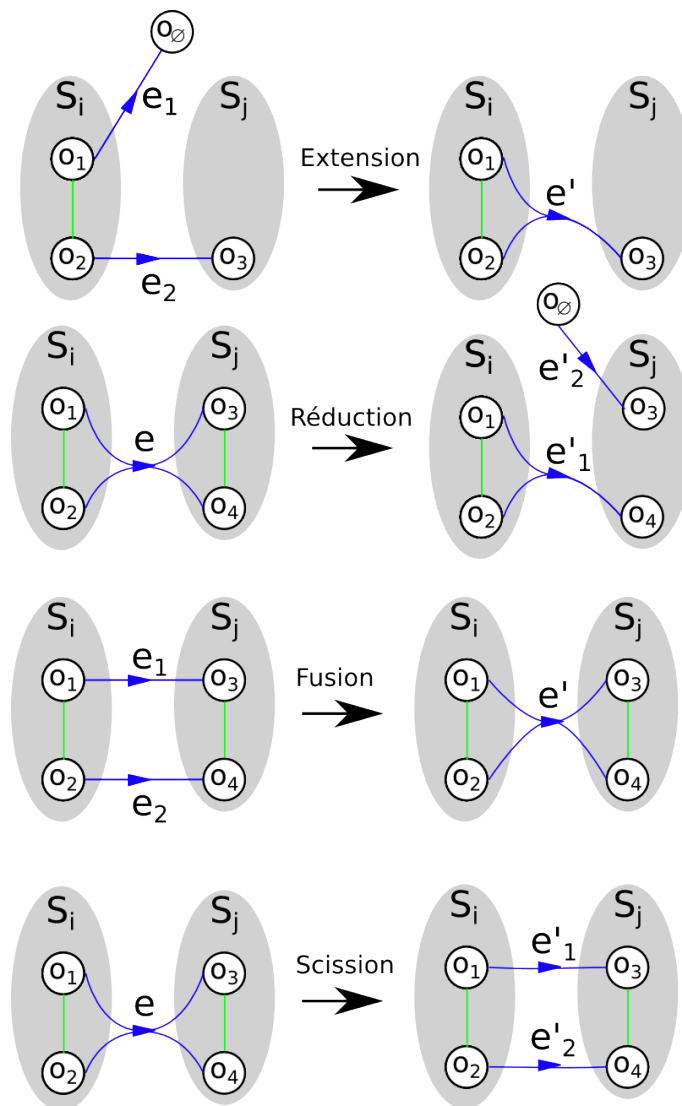


FIGURE 6.8 – Illustration des modifications d’extension, de réduction, de fusion et de scission.



FIGURE 6.9 – Schéma général de la création d’une modification d’un hypergraphe de filiation.

ou une disparition et lie donc une observation à l'observation nulle. Dans le second cas, il s'agit d'un cas particulier de l'extension qui revient à créer un nouvel appariement entre deux observations que l'on désigne comme une modifications spécifique de **naissance** d'un appariement. Dans ce cas précis, il est nécessaire de retirer l'observation nulle de l'hyperarc résultant de la fusion.

L'étape de filtrage dépend du cas de figure :

1. S'il s'agit d'une naissance, l'hyperarc à étendre est de la forme $e = (\{o_1\}, \{o_\emptyset\})$ ou $e = (\{o_\emptyset\}, \{o_1\})$. Si e est de la première forme, les hyperarcs candidats sont tous ses voisins de la forme $e_2 = (\{o_\emptyset\}, \{o_2\})$. La situation est symétrique si e est de la seconde forme. Les hyperarcs voisins doivent ensuite être filtrés pour ne conserver que ceux qui permettent de créer un nouvel arc $e' = (\{o_1\}, \{o_2\})$ ou $e' = (\{o_2\}, \{o_1\})$ admissible. La contrainte C_2 est d'ors et déjà vérifiée pour tous les arcs voisins. La contrainte C_1 peut être aisément vérifiée en calculant l'antécédence de o_1 et o_2 . Pour vérifier C_3 , il faut vérifier qu'il n'existe pas de chemin dans le graphe allant de l'observation constituant la tête de e' vers celle constituant sa queue. Un exemple de filtrage dans le cas d'une naissance est illustré en figure 6.10a. Les observations candidates valides pour l'hyperarc e sont indiquées par un contour rose pointillé, celles invalides étant en rouge. Les relations spatiales existantes sont indiquées par des liens en vert.
2. Si e est de cardinalité supérieure ou égale à 1 : 1, les arcs candidats sont tous ses voisins de cardinalité 1 : 0 ou 0 : 1. Le filtrage des candidats valides peut être effectué de la même façon que pour la naissance. Un exemple de filtrage dans le cas d'une extension est illustré en figure 6.10b. La légende est similaire à l'exemple donné pour la naissance.

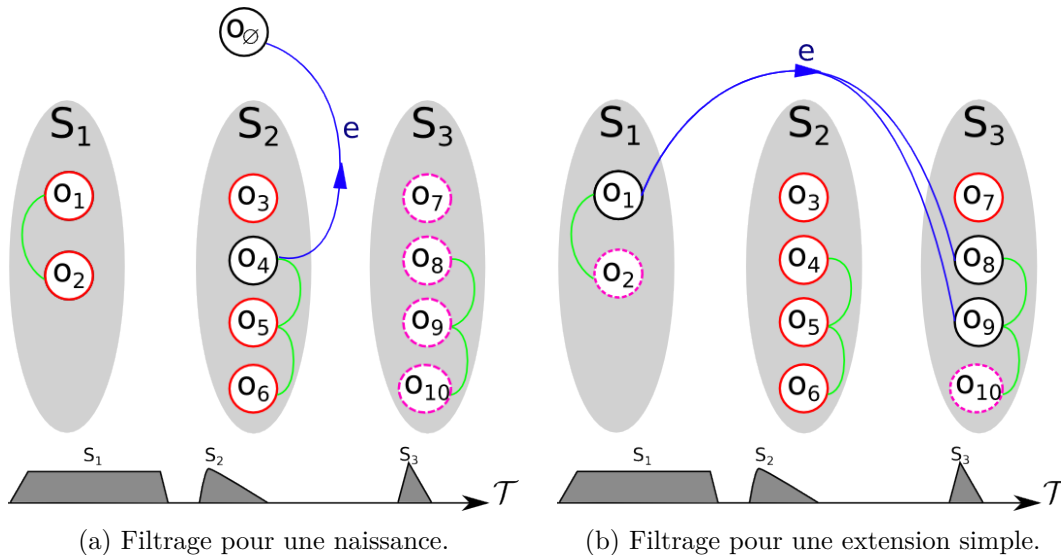


FIGURE 6.10 – Exemples de l'étape de filtrage pour une modification de type "extension". Seul l'hyperarc concerné par la modification est montré.

Une fois le filtrage effectué, la modification peut être créée en sélectionnant de manière aléatoire un arc voisin candidat et valide e_i puis en fusionnant e et e_i .

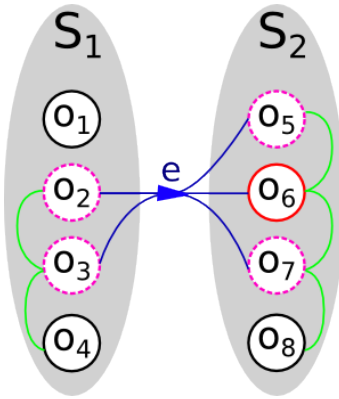


FIGURE 6.11 – Exemples de l'étape de filtrage pour une modification de type "réduction".

Avec $e = (Q, T)$ et $e_i = (Q_i, T_i)$, cette fusion consiste à créer un nouvel hyperarc $e' = (Q \cup Q_i \setminus \{o_\emptyset\}, T \cup T_i \setminus \{o_\emptyset\})$.

Réduction d'un hyperarc

La réduction d'un hyperarc e est une modification symétrique de l'extension. Là où l'extension permettait d'ajouter une observation à un arc existant, la réduction permet de retirer une observation de e . Ceci revient à diviser e en deux nouveaux arcs e'_1 et e'_2 , le premier étant de cardinalité $0 : 1 / 1 : 0$ et le second d'une cardinalité supérieure ou égale à $0 : 1 / 1 : 0$. De la même façon que l'extension pouvait aboutir à une naissance, la réduction peut créer une **mort**, c'est à dire à division d'un arc en deux nouveaux arcs correspondant à une apparition et une disparition.

Contrairement à l'extension, la réduction opère sur les observations composant l'arc. Les candidats ne sont donc pas des arcs voisins de e mais des arcs qu'il est possible de créer à partir des observations membres de sa queue et de sa tête, ce qui revient plus simplement à choisir une observation de e à lui retirer. Le fait de gérer l'extension comme une agrégation d'arcs voisins permet d'obtenir des hyperarcs dont les sous-graphe du graphe des relations spatiales G_s induits par les observations composant sa tête et sa queue ne possèdent qu'une seule composante connexe¹⁷. Pour conserver la symétrie avec la modification d'extension, la réduction doit préserver cette propriété. Pour cela, le filtrage doit éliminer les observations de l'hyperarc qui ne sont pas dans la frontière intérieure¹⁸ des sous-graphes spatiaux induits¹⁹ par sa queue et sa tête. Ainsi, nous assurons de toujours conserver une seule composante connexe. Enfin, nous interdisons le fait de retirer une observation d'un hyperarc décrivant une apparition ou une disparition, afin de ne pas créer d'hyperarcs du type $e = (\{o_\emptyset\}, \{o_\emptyset\})$. La figure 6.11 illustre le filtrage des candidats dans le cas d'une réduction. Tout comme l'extension, la réduction peut être créée après filtrage par sélection aléatoire d'une observation candidate et valide o_i puis en divisant e

17. Ceci traduit une hypothèse implicite sur les processus spatio-temporels : ils concernent des observations liées spatialement.

18. La frontière intérieure des sommets du sous-graphe W d'un graphe V désigne tous les sommets de W étant connectés à au moins un sommet de $V \setminus W$.

19. Un sous-graphe induit W d'un graphe V est un sous-graphe défini par un ensemble de sommets. Les arêtes de W correspondent aux arêtes de V incidentes à deux sommets de W .

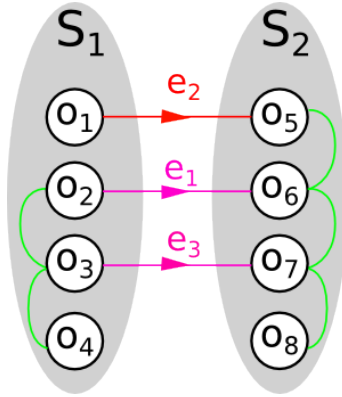


FIGURE 6.12 – Exemples de l'étape de filtrage pour une modification de type "fusion".

en deux nouveaux arcs. Si e est de la forme $e = (Q, T)$ et si $o_i \in Q$, les deux arcs créés sont $e_1 = (\{o_i\}, \{o_\emptyset\})$ et $e_2 = (Q \setminus \{o_i\}, T)$. Si $o_i \in T$, $e_1 = (\{o_\emptyset\}, \{o_i\})$ et $e_2 = (Q, T \setminus \{o_i\})$.

Fusion de deux hyperarcs

Créer une nouvelle relation de filiation entre deux observations de sources différentes peut aboutir à la fusion de deux hyperarcs. La modification de fusion consiste à agréger deux hyperarcs voisins e_1 et e_2 de cardinalités supérieures ou égales à 1 : 1. Il en résulte un nouvel arc e' de cardinalité $n : m$, avec $n = |Q(e_1)| + |Q(e_2)|$ et $m = |T(e_1)| + |T(e_2)|$.

L'étape de filtrage consiste à éliminer les arcs candidats ne respectant pas les contraintes C_2 ou C_3 . La vérification du respect de ces deux contraintes est toutefois plus complexe que dans les cas précédents. Pour C_3 , il est nécessaire de vérifier qu'il n'existe aucun chemin menant de la tête de e_2 vers la queue de e_1 ou de la tête de e_1 vers la queue de e_2 . Il y a donc au plus $n + m$ explorations à effectuer. Toutefois, sauf dans le cas de graphes dont les observations présentent des transformations particulièrement complexes, le nombre de chemins existant à partir d'une observation est relativement faible. Pour C_2 , une méthode simple consiste à calculer l'intersection des temporalités des observations de $Q(e_1) \cup Q(e_2)$ puis celle des temporalités de $T(e_1) \cup T(e_2)$. Si l'une des deux est vide, la contrainte n'est pas respectée.

La figure 6.12 présente un exemple de filtrage d'une modification de fusion. Les arcs e_1 et e_2 ne peuvent être fusionnés car ils ne sont pas voisins. Les arcs e_1 et e_3 sont pour leur part des candidats valides être fusionnés. À partir d'un hyperarc e_1 fixé, la création d'une modification de fusion est effectuée en sélectionnant un hypergraphe voisin valide de e_2 puis en créant un nouvel hyperarc e' dont la tête et la queue sont formées de l'union des têtes et queues de e_1 et e_2 .

Cet arc est de la forme $e' = (Q(e_1) \cup Q(e_2), T(e_1) \cup T(e_2))$

Scission d'un hyperarc

La dernière modification possible est la scission d'un hyperarc e en deux autres hyperarcs de cardinalité supérieure ou égale à 1 : 1. Puisque nous nous assurons de toujours manipuler des hypergraphes admissibles, aucune contrainte ne doit être vérifiée pour diviser un arc. Cependant, si l'on veut obtenir deux nouveaux arcs

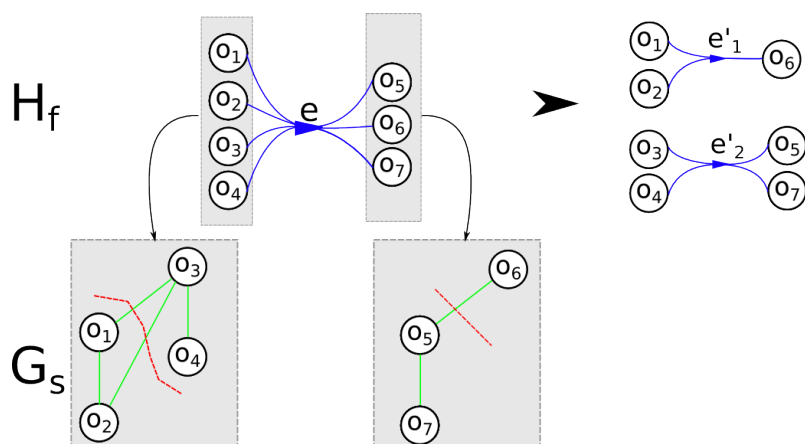


FIGURE 6.13 – Exemples de la scission d’un hyperarc en utilisant une coupe aléatoire dans les sous-graphes spatiaux induits par sa tête et sa queue.

n’étant pas des apparitions ou des disparitions, l’arc e doit nécessairement être de cardinalité $n : m$. Pour la scission, l’étape de filtrage consiste donc simplement à vérifier la cardinalité de e : si elle ne convient pas, l’arc ne peut être scindé.

La génération des candidats pour une scission est particulière puisque cette modification consiste à diviser un arc en deux nouveaux arcs de cardinalité moindre. Il faut donc scinder la queue de e en deux sous-ensembles non vides Q_1 et Q_2 puis faire de même pour sa tête pour obtenir T_1 et T_2 . Deux nouveaux arcs e'_1 et e'_2 peuvent être alors créés à partir de ces quatre sous ensembles en associant ceux-ci deux à deux. Deux problèmes se posent cependant. D’une part, nous souhaitons conserver des hyperarcs dont les observations sont connectées spatialement de façon à former de nouveaux arcs cohérents avec la façon de traiter les autres modifications. De plus, plusieurs couples d’arcs peuvent être formés une fois la tête et la queue scindée. Il faut donc effectuer un choix sur la façon d’associer ces sous-ensembles en nouveaux arcs.

Un sous-ensemble d’observations respecte la contrainte de connexion spatiale si le sous-graphe spatial induit par ces observations est connexe. Diviser un ensemble d’observations en deux sous-ensembles dont les éléments sont connectés revient donc à effectuer une coupe dans le sous-graphe spatial induit de façon à obtenir exactement deux composantes connexes. Une façon naïve de réaliser cette opération consiste à retirer du sous-graphe des arcs jusqu’à obtenir deux composantes connexes²⁰. De cette façon, nous obtenons des sous-ensembles Q_1, Q_2, T_1 et T_2 dont les sous-graphes spatiaux induits sont connexes. Afin d’associer les ensembles d’observations ainsi créés en nouveaux arcs, nous proposons de créer aléatoirement deux couples (Q, T) , ce qui permet de créer une scission aléatoire d’un hyperarc. La figure 6.13 présente un exemple du découpage d’un hyperarc e en deux nouveaux arcs en effectuant une coupe aléatoire des sous-graphes spatiaux induits par $Q(e)$ et $T(e)$. Une autre possibilité serait d’associer les ensembles Q_i et T_j dont les observations sont les plus proches spatialement. Cependant, le comportement stochastique de la modification serait partiellement perdu.

20. Il s’agit de fait d’une coupe aléatoire.

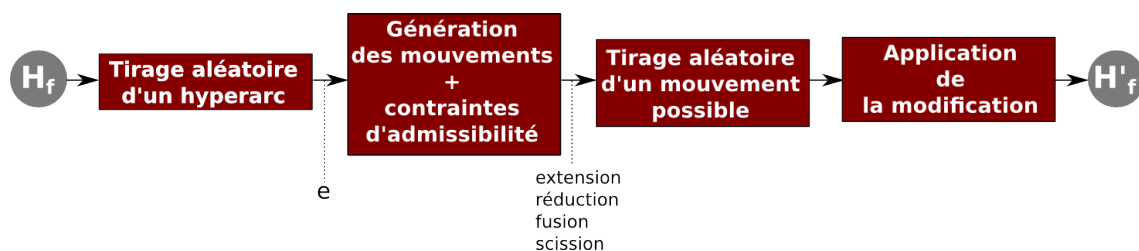


FIGURE 6.14 – Génération et sélection d’une solution du voisinage de la solution courante.

Génération d’un hypergraphe voisin du graphe courant

Dans les paragraphes précédents nous avons expliqué comment pouvaient être créées des modifications sur un hyperarc donné d’un hypergraphe de filiation de façon à générer un nouvel hypergraphe admissible. Le voisinage d’un hypergraphe est donc formé de tous les hypergraphes admissibles résultants de l’application d’une modification à l’un de ses hyperarcs. Au cours du recuit simulé, il n’est pas nécessaire de produire l’ensemble du voisinage de l’hypergraphe courant²¹ mais seulement de tirer aléatoirement un hypergraphe voisin qui sera évalué par le critère de Metropolis. La manière de sélectionner un hypergraphe voisin d’un hypergraphe donné est illustré dans le schéma 6.14.

Évaluation de la solution voisine choisie

L’étape d’évaluation d’un hypergraphe H'_f sélectionné aléatoirement comme nouvelle solution du recuit simulé consiste à calculer la fonction $Iv(H'_f)$. Ce calcul a été détaillé dans la section précédente et nécessite d’agrèger les observations constituant les extrémités des arcs de H'_f . Le critère de Metropolis peut ensuite être appliqué en comparant la valeur de $Iv(H'_f)$ avec $Iv(H_f)$, H_f étant le graphe courant. La probabilité d’accepter H'_f comme solution courante de l’algorithme est donc :

$$p(\Delta Iv) = e^{\left(-\frac{Iv(H'_f) - Iv(H_f)}{T}\right)} \quad (6.11)$$

Loi de décroissance de la température

La loi de décroissance de la température du recuit $f(T)$ est un élément crucial de la conception d’un algorithme efficace. Elle doit permettre une décroissance de la température au fil du temps suffisamment lente pour permettre à l’algorithme de converger vers une solution proche de l’optimum, tout en permettant cette convergence dans un temps raisonnable. Le choix d’une loi de décroissance repose sur un compromis entre qualité de la solution finale et rapidité de l’algorithme et dépend donc du type de problème, de sa complexité et des contraintes sur son temps de réponse. Dans notre cas, le temps de calcul nécessaire pour la découverte des relations de filiation n’est pas un critère bloquant puisqu’il peut être effectué une seule fois pour plusieurs interprétations de l’identité. De plus, nous visons des solutions proches de l’optimum pour réduire les corrections manuelles et, surtout, éviter les cas de sous-appariement. Nous choisissons pour cela une loi de décroissance géométrique

21. Ce qui ne serait de toute façon pas raisonnable en termes de temps de calcul.

proposée par Kirkpatrick *et al.* (1983), de la forme :

$$f(T) = T_0\mu^i \quad (6.12)$$

où i est l'itération courante et μ est un facteur de décroissance dont la valeur est définie entre 0 et 1. Il a été montré expérimentalement que la décroissance géométrique a un comportement optimal si μ est proche de 1 (Kirkpatrick *et al.*, 1983). Dans un processus de recuit simulé classique, la température ne décroît que par paliers et reste fixe pour un certain nombre d'itérations. Ce nombre d'itérations étant un paramètre empirique, Lundy et Mees (1986) montrent qu'il peut être supprimé à condition de faire décroître la température lentement à chaque itération. Notons enfin que le nombre d'itérations à chaque palier de température peut être modifié durant le déroulement de l'algorithme. L'idée est alors d'effectuer de descendre rapidement la température lorsque celle-ci est élevée, puis d'allonger petit à petit la durée de chaque palier au fur et à mesure que la température diminue. De cette façon, le parcours de l'espace de recherche est dans un premier temps proche d'une marche aléatoire, puis le rayon de recherche va petit-à-petit se resserrer autour d'une zone contenant de bonnes solutions.

Température initiale T_0

La température initiale d'un recuit simulé est également un paramètre critique qui conditionne la convergence vers des bonnes solutions. Si celle-ci est trop faible, le risque est d'être rapidement bloqué dans un minimum local de l'espace de recherche. Trop élevée, elle empêchera l'algorithme de converger. Ce paramètre peut être fixé de manière empirique par expérimentation, où être calculé a priori lors d'une phase consistant à explorer l'espace de recherche afin de déterminer une température initiale assurant un certain taux d'acceptation de mauvaises solutions. Différents travaux se sont penchés sur cette questions. Rayward-Smith *et al.* (1996) proposent ainsi d'effectuer un premier recuit avec une température initiale très élevée et décroissant rapidement jusqu'à ce qu'environ 60% de solutions dégradant la fonction objectif soient acceptés. La température correspondant à cet instant est ensuite utilisée comme température initiale d'un second recuit plus lent. Dowsland et Thompson (2012) proposent un processus similaire consistant à effectuer un premier recuit en augmentant rapidement la température jusqu'à ce qu'une certaine proportion de mauvaises solutions soient acceptées, puis de décroître lentement la température afin de trouver la solution optimale. Citons également Ben-Ameur (2004) qui introduit une méthode permettant d'estimer une température initiale assurant un taux d'acceptation de 80% d'une solution de l'espace de recherche. Comme les autres, cette méthode consiste à parcourir l'espace de recherche à haute température, puis à faire décroître celle-ci jusqu'à obtenir un taux d'acceptation moyen de 0.8. Dans notre cas, nous choisissons de nous appuyer sur la proposition de Ben-Ameur (2004) afin de fixer la température initiale.

Génération du graphe initial H_{f_0}

La solution initiale utilisée pour amorcer l'optimisation peut être fixée -par exemple un graphe vide- ou choisie de manière aléatoire. Cette seconde possibilité assure une meilleure couverture de l'espace de recherche sur plusieurs lancements de

l'algorithme. Pour générer un hypergraphe de filiation initiale aléatoire, une solution est d'effectuer une marche aléatoire dans l'espace de recherche en réalisant une première phase de recuit à haute température. En particulier, la solution initiale peut être celle créée par l'étape d'évaluation de la température initiale.

Conclusion sur la découverte des relations de filiation

Nous avons proposé dans cette section une méthode automatique s'appuyant sur un recuit simulé mono-objectif afin de réaliser la phase de découverte des relations de filiations. Cette phase est une étape d'appariement de données géographiques qui s'appuie sur une connaissance des processus spatio-temporels entre observations géo-historiques et sont modélisés à l'aide d'un hypergraphe de filiations²². La résolution de cette phase d'appariement par un processus de recuit simulé nous permet de créer un hypergraphe finale H_f optimal ou proche de l'hypergraphe optimal H^*_f . Cependant, l'hypergraphe produit ne décrit pas réellement les processus spatio-temporels à l'œuvre car il n'a pas connaissance du type des relations de filiations. En effet, la phase d'appariement permet de détecter les observations liées par une relation de filiation mais elle ne détermine pas si ces relations sont des continuations ou des dérivations. Dans ce but, nous introduisons une étape de typage des relations de filiations qui ont été identifiées lors de la phase d'appariement.

6.3 Typage des relations de filiations

Une fois l'étape de découverte des relations de filiation effectuée, nous disposons d'un hypergraphe représentant des groupes d'observations affiliées. Le type de ces relations de filiations dépend d'une définition précise de l'identité des observations qui n'a pas été nécessairement définie au moment de l'appariement. Ceci permet, à partir d'un seul appariement, de produire plusieurs graphes géohistoriques selon une certaine vision de l'identité des observations par un utilisateur. Dans cette section, nous présentons une approche permettant d'instancier un graphe de filiations à partir d'un hypergraphe de filiation à l'aide d'une comparaison entre les identités des observations qui ont été liées. De façon à être cohérente avec la phase d'appariement et pour prendre en compte les incertitudes liées aux observations, cette comparaison s'appuie également sur la théorie des croyances.

Le typage des relations de filiations découvertes est un processus en 3 étapes, résumé par le schéma 6.15. Nous présentons dans cette section les différentes étapes de ce processus. Dans un premier temps, nous présentons la création d'un graphe de filiation non typé G à partir d'un hypergraphe de filiations. Ensuite, nous proposons une approche permettant d'étiqueter chaque arc de ce graphe selon le type de filiation qu'elle représente à l'aide de la théorie des croyances et d'un ensemble de mesures sur l'identité des observations. Enfin, une dernière étape permet de résoudre les conflits générés par cette phase de typage de façon à créer un graphe de filiation valide. Une fois ces étapes réalisées, nous disposons d'un graphe géohistorique

22. Les méthodes d'appariement de données géographiques s'appuient sur différentes mesures de similarité entre objets. Celles-ci dépendant des données, elles seront abordées ultérieurement lors de l'application de l'approche aux observations sur Paris. Seules les mesures portant sur le temps sont génériques et sont données dans la section suivante pour tout type d'objet géographique.

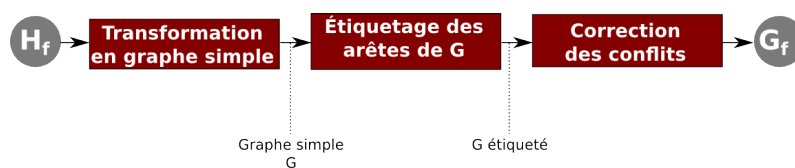


FIGURE 6.15 – Schéma général de la phase de typage des filiations.

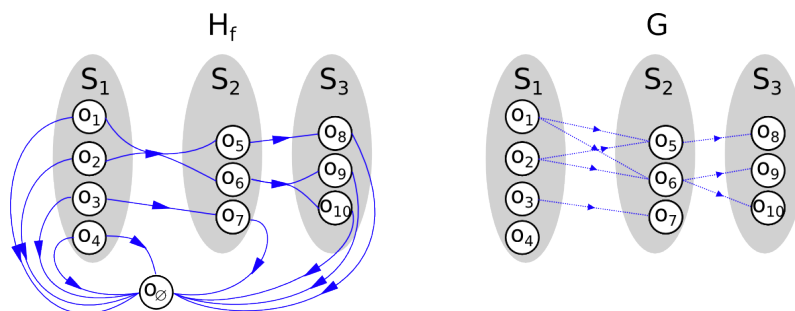


FIGURE 6.16 – Décomposition d'un hypergraphe de filiation en graphe simple.

complet compatible avec le modèle de Del Mondo (2011).

6.3.1 Décomposition d'un hypergraphe H_f

La première étape du processus de typage des relations de filiations consister à décomposer l'hypergraphe de filiation construit $H_f = (O, E_H)$ par optimisation en un graphe orienté simple $G(O, A)$ composé d'arcs correspondant à des relations de filiations entre observations.

Cette décomposition, illustrée en figure 6.16, peut être effectuée en deux étapes :

1. Un graphe vide $G(O, \emptyset)$ est créé.
2. Pour tout hyperarc $e = (Q, T)$ de H_f , le produit cartésien $Q \times T$ produit $|Q| \cdot |T|$ couples d'observations. Pour chacun de ces couples est ajouté dans G un arc liant les deux observations membres du couple.

Ainsi, nous obtenons un premier graphe G non étiqueté qui peut servir de support à la création d'un graphe de filiation.

6.3.2 Étiquetage des filiations par fusion d'informations incertaines

Le graphe G produit décrit une ensemble relations de filiations dont la nature doit être précisée. Cette nature dépend de l'identité des observation -et donc de leur nature-, mais également d'une certaine vision de la transmission de l'identité par un utilisateur. Si l'on prend le cas de rues, un utilisateur s'intéressant à l'évolution de leur nommage considérera le nom comme membre de l'identité, un changement de nom entre deux rues résultera en une dérivation. Si seule la morphogenèse du réseau viaire est étudiée, un changement de nom n'aura aucun impact, la filiation résultante sera une continuation. Dans les deux cas, le graphe G support est le même, seule la définition de l'identité change.

Enfin, le graphe créé par décomposition de H_f contient un certain nombre d'arcs

qu'il n'est pas toujours souhaitable de conserver. Le graphe G est une structure générale qui doit être raffinée. Ceci est particulièrement vrai pour les hyperarcs de cardinalité $n : m$ dont la décomposition crée artificiellement des arcs n'ayant pas toujours de sens. Enfin, le processus d'appariement par optimisation peut aboutir à des situations de sur-appariement, en particulier avec des arcs $n : m$ lorsque la transformation des observations est ambiguë (en raison des imprécisions, du manque de connaissances, etc.). Nous introduisons pour ces deux raisons la possibilité d'étiqueter une arc de G comme une relation erronée devant être supprimée du graphe de filiation final.

L'étape de découverte des relations de filiations s'appuie sur la théorie des croyances afin d'identifier les observations liées. Cette théorie offrant l'avantage de modéliser l'incertitude et l'imprécision des informations, elles est également pertinente pour étiqueter les arcs de G . Dans cette section, nous présentons une procédure d'étiquetage des filiations utilisant la théorie des croyances.

Fonction d'étiquetage

Nous définissons trois étiquettes $\{\delta, \gamma, \epsilon\}$ pouvant être assignées à toute arc du graphe G , avec la signification suivante :

- Une arête portant l'étiquette δ est relation de type dérivation,
- Une arête portant l'étiquette γ est relation de continuation,
- Une arête portant l'étiquette ϵ est relation erronée en représentant pas de réelle filiation.

Nous introduisons une fonction d'étiquetage d'un arc $a \in A$ de $G = (O, A)$

$$r : A \rightarrow \{\gamma, \delta, \epsilon\} \quad (6.13)$$

Nous devons maintenant détailler cette fonction afin de pouvoir étiqueter tout graphe G construit à partir d'un hypergraphe de filiation.

Sources d'informations pour l'étiquetage

Tout comme pour l'étape de découverte, l'étiquetage repose donc sur des mesures de similarité d'identité. Ces mesures sont spécifiques à la nature des observations et donc au cas applicatif. Celles-ci peuvent être différentes des mesures utilisées dans l'étape de découverte des filiations. En effet, si elles comparent également des éléments de l'identité de deux observations, elles visent un objectif plus précis que lors de l'appariement. Il est cependant nécessaire que tous les éléments de l'identité soient comparés pour que le typage des arcs soit cohérent avec la définition de l'identité des observations spécifiée en amont de la découverte des relations de filiation.

Cadre de discernement

Identifier la nature exacte d'une relation de filiation revient à comparer les identités des observations de l'arc de G décrivant cette relation. Tout comme pour la phase de découverte, cette comparaison peut s'appuyer sur un ensemble de mesures entre les membres de l'identité des observations en jeu. Nous avons alors pour chaque arc de G un problème de décision multicritère, que nous souhaitons résoudre à l'aide de la théorie des croyances. Notons que nous sommes cette fois dans un cas de monde fermé, la décision du type de chaque arc ne dépendant que des observations qui

constituent ses extrémités. L'étiquetage étant de la même forme pour tous les arcs du graphe, nous pouvons réduire le problème de décision à un arc $a \in A$ seul.

Nous définissons pour a un cadre de discernement $\Omega = \{\gamma, \delta, \epsilon\}$ correspondant aux étiquettes qui peuvent lui être assignées. Le problème est donc de choisir une étiquette maximisant la croyance d'un ensemble de sources d'informations sur le type de a .

Masses de croyance

Nous ne définissons cette fois pas de modèle spécifique pour les fonctions de masse. Ainsi, il est possible que des sources d'information sur l'identité placent une croyance sur des disjonctions de propositions dans Ω . Considérons par exemple le cas de deux parcelles cadastrales dont l'identité est formée de leur géométrie et de leur numéro, liées par une relation de filiation dont le type est inconnu. Un changement de numéro de parcelle ne signifie pas nécessairement que la relation est une dérivation : cela dépend en réalité de ce que l'on étudie. Par contre, un numéro identique est un indice tendant à dire qu'il s'agit de la même parcelle et que la relation est une continuation. Dans ce cas, une source d'information pourrait avoir le comportement suivant :

- si les numéros sont identiques, la source croit fortement en la proposition $\{\gamma\}$.
- si les numéros sont différents, la source est incapable de dire s'il s'agit d'une continuation ou d'une dérivation. Sa croyance serait alors placée sur la disjonction $\{\gamma, \delta\}$.

Comme les fonctions de masse pour le typage ne suivent aucun modèle précis, leur définition dépend totalement du type d'observation et de la nature de la source d'information. Ces fonctions de masse doivent donc être définies pour une application précise. Ainsi, les fonctions de masse pour le typage des relations dans le cas des rues de Paris sont présentées dans la section 6.4.5.

Fusion des masses et décision

Comme nous l'avons dit, nous sommes cette fois dans le cas d'un monde fermé. La fusion de l'ensemble des masses de croyances définies sur les éléments de 2^Ω doit être effectuée à l'aide de l'opérateur conjonctif de Dempster défini précédemment. La dernière étape après fusion des croyances consiste à décider de l'étiquette assigner à l'arc, et donc du type de relation. Afin de prendre en compte les masses de croyance placées sur les disjonctions de propositions, nous proposons d'utiliser le principe du maximum de probabilité pignistique. Ce mode de prise de décision consiste à choisir la proposition de Ω dont la probabilité pignistique est la plus grande.

6.3.3 Construction d'un graphe géohistorique valide

L'étiquetage des arcs de G ne suffit pas pour produire un graphe de filiations G_f valide. En effet, chaque arc ayant été considéré comme un problème de fusion d'information indépendant, nous pouvons obtenir des graphes dans lesquels une observation est en relation de continuation avec plusieurs autres observations. Ceci a pourtant été interdit (voir le chapitre 5) pour conserver une cohérence du graphe vis-à-vis de la notion d'identité. En effet, un tel cas de figure signifierait que plusieurs

observations d'une même sources géohistorique partagent la même identité. Nous proposons ici une étape de post-traitement visant à résoudre ce type de conflit.

Résolution des conflits liés aux continuations multiples

Une façon simple de résoudre les conflits dans le graphe créé consiste à identifier les sous-ensembles d'arcs de G en conflit. Pour chaque sous-ensemble, il ne faut conserver qu'un seul arc étiqueté comme continuation. Pour déterminer cet arc, il suffit de ré-appliquer le principe de maximum de probabilité pignistique. Tous les autres arcs sont alors étiquetés comme dérivations.

Niveau de détail

Cette méthode simple est toutefois naïve car elle ne permet pas de détecter les situations où les conflits correspondent en réalité à des changements de niveau de détail entre deux sources géohistoriques. Nous n'avons toutefois pas exploré plus en profondeur la phase de résolution des conflits afin de corriger ce type de cas. Une possibilité serait d'analyser les écarts entre probabilités pignistiques des arcs en conflit et, en dessous d'un certain seuil, considérer qu'il s'agit d'un cas de changement de niveau de détail.

6.4 Mise en œuvre et résultats : application aux rues de Paris

Dans cette section, nous appliquons l'approche de construction de graphe géohistoriques aux réseaux des rues de Paris extrait des différents atlas parisiens traités dans les chapitres précédents.

6.4.1 Implémentation de l'approche de construction de graphes géohistoriques

Un prototype expérimental de l'approche proposée a été implémentée au sein de la plateforme SIG libre Geoxygene²³. Cette plateforme est développée en Java au sein du laboratoire COGIT de l'Institut National de l'Information Géographique et Forestière²⁴. Nous avons mis en place les deux modules principaux de découverte des relations de filiation et de création d'un graphe de filiation par typage. Ces deux modules interagissent avec la base de données spatiale et temporelle stockant les *snapshots* sur les rues de Paris. De plus, nous stockons les graphes créés à l'issue des deux modules également dans une base de données relationnelle. Ces deux bases de données ont été mises en place à l'aide du SGBD Postgresql muni de l'extension Postgis pour les données géographiques et du module développé par Van Daele (2014) pour la représentation de sous-ensembles flous.

Notre implémentation utilise de plus certaines bibliothèques externes à Geoxygene afin d'intégrer la théorie des croyances, de manipuler efficacement des hypergraphes et enfin de manipuler des sous-ensembles flous. Concernant la théorie des croyances,

23. <http://oxygene-project.sourceforge.net/>

24. <http://recherche.ign.fr/labs/cogit/accueilCOGIT.php>

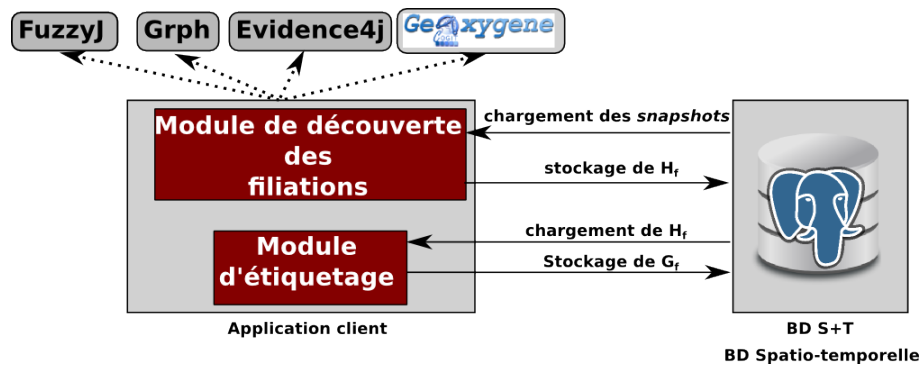


FIGURE 6.17 – Schéma général du prototype implémenté.

nous nous appuyons sur la bibliothèque libre Java Evidence4j^{25 26} intégrant toutes les fonctionnalités nécessaires. La modélisation de l'hypergraphe de filiation et des graphes créés lors de l'étape de typage s'appuient sur la bibliothèque opensource Grph²⁷ développée au laboratoire I3S de l'université de Sophia Antipolis. Enfin, pour manipuler des sous-ensembles temporels flous, nous avons utilisé la bibliothèque également libre FuzzyJ²⁸ développée par le *National Research Council Canada*. L'architecture générale du prototype développé est illustré en figure 6.17. Rajoutons également que le code de ce prototype devrait être déposé en OpenSource rapidement.

6.4.2 Données utilisées

Les quatre atlas dont les rues ont été vectorisés sont les suivants :

- l'atlas de Verniquet, dont le temps valide est le sous-ensemble flou (1783, 1785, 1791, 1799),
- l'atlas par îlots de Vasserot, entièrement vectorisé par Noizet *et al.* (2008), dont nous avons fixé le temps valide (1808, 1810, 1836, 1854),
- l'atlas de Jacoubet, de temps valide (1825, 1827, 1836, 1837),
- l'atlas municipal de Paris de l'année 1888, de temps valide (1887, 1888, 1889),

Le temps valide désigne la période pendant laquelle les observations des entités du monde réel sont effectuées puis transcrites dans l'atlas. Par exemple, nous considérons que l'atlas de Verniquet décrit des entités du monde réel ayant existé entre 1783 et 1799. La certitude de cette hypothèse n'est pas toujours totale, ce qui est exprimé par l'utilisation des sous-ensembles temporels flous. Ainsi, nous sommes sûrs que l'atlas de Verniquet ait pu permettre l'observation d'entités du monde réel entre 1785 et 1791 c'est à dire la période de levé topographique et de dessin de l'atlas. Avant 1785 et après 1791, il est de moins en moins certain que des relevés ou des corrections aient été effectuées.

La figure 6.18 illustre la localisation temporelle de ces différents atlas. On peut remarquer que le temps valide de l'atlas de Jacoubet est totalement inclus dans

25. <https://github.com/IGNF/evidence4j>

26. Elle même fondée sur la bibliothèque C++ Evidenz(Burrus et Lesage, 2007).

27. <http://www.i3s.unice.fr/~hogie/grph/>

28. <http://rorchard.github.io/FuzzyJ/>

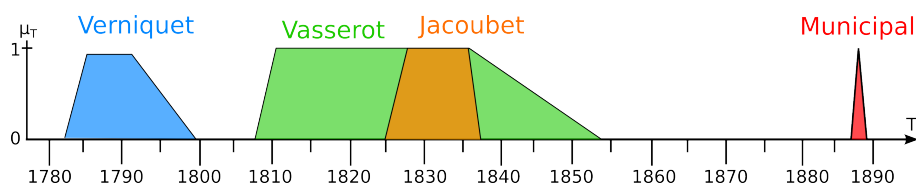


FIGURE 6.18 – Temps valides des différents atlas.

celui de l’atlas de Vasserot. Cela signifie que la méthode de construction du graphe considérera comme possible l’existence de relations de filiations allant de Vasserot à Jacoubet ou de Jacoubet vers Vasserot.

Nous n’avons pas l’information détaillée des temps valide pour chaque observations vectorisée des rues de Paris. Bien que ces informations puissent être extraites de sources telles que les dictionnaires de rues (voir (Lazare et Lazare, 1844, 1855)), leur saisie dépasse de la cadre de cette thèse. Pour cette raison, **nous avons considéré que les observations des rues de Paris ont le même temps valide que leur source géohistorique.**

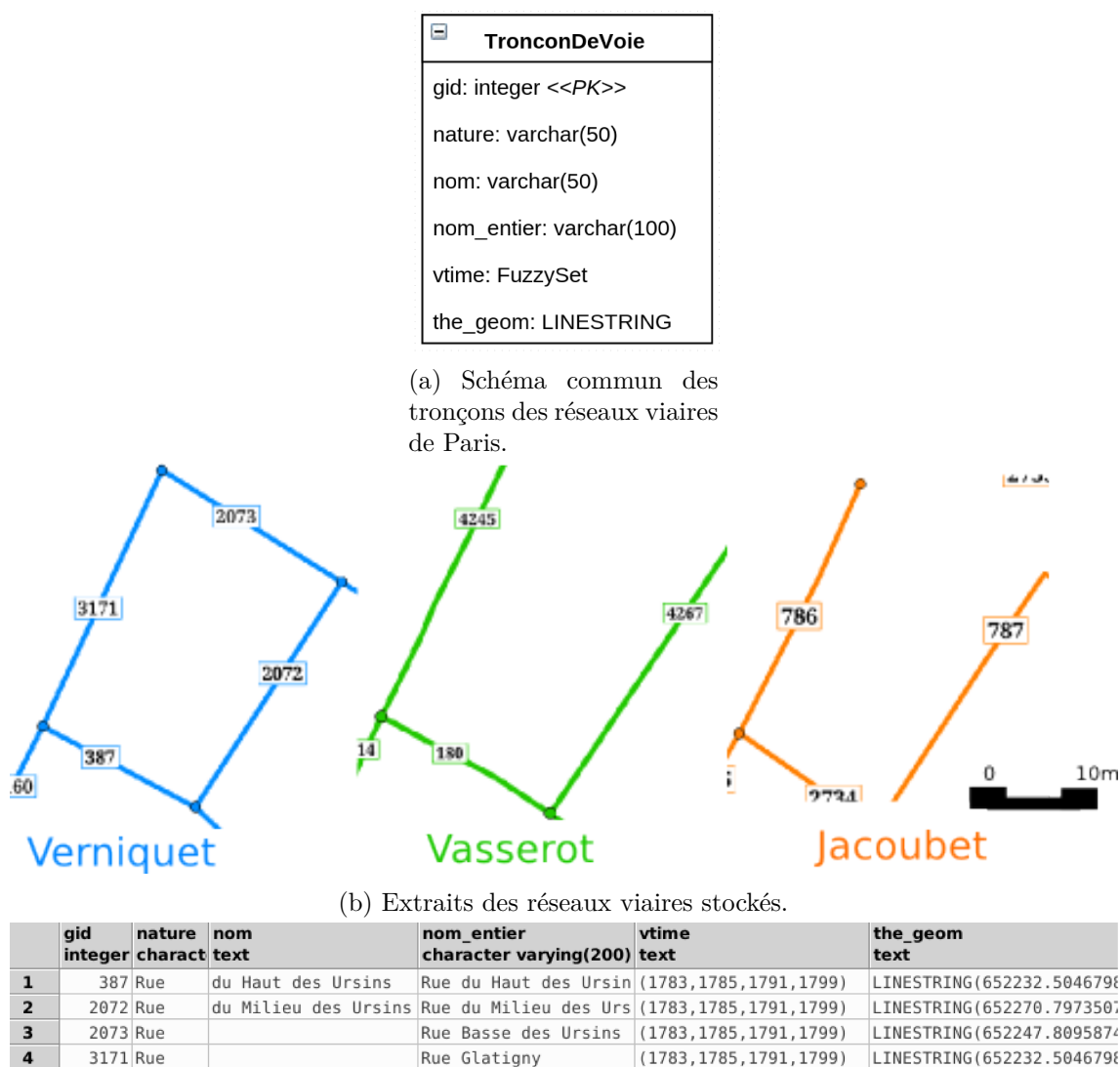
Représentation des réseaux viaires dans la base de données spatiale et temporelle

Les réseaux de rues vectorisés et présentés en fin du chapitre 3 sont des graphes planaires topologiquement justes. Chaque réseau de rue vectorisé a été stocké comme un *snapshot* vecteur de notre base de données géohistorique. Les observations géohistoriques de ces *snapshots* sont des tronçons de rues, et non des rues au sens commun du terme (c’est-à-dire un ensemble contigu de tronçons portant le même nom). Dans la base de données spatiale et temporelle, ces observations sont modélisées sous la forme de *feature* linéaires²⁹. La nature de ces observations est importante à garder en tête, en particulier pour l’étape de typage des filiations. En effet, ce ne seront alors pas les relations de filiations entre rues qui seront identifiées, mais les relations de filiation entre tronçons. Les transformations de l’espace qui en résultent sont donc également entre tronçons. La figure 6.19 illustre ainsi le découpage du réseau viaire en tronçons de rues représentés sous la forme de *features* linéaires. Cette figure présente tout d’abord le schéma de donnée aligné des tronçons de rue issus des différents *snapshots* que nous utilisons pour tester notre méthode (6.19a), puis illustre le découpage du réseau en tronçons pour les *snapshots* des tronçons de Verniquet, Vasserot et Jacoubet (6.19b). Enfin, un extrait du snapshot des tronçons de l’atlas de Verniquet est présenté en figure 6.19c. Les champs *vtime* (temps valide) et *the_geom* (géométrie) sont figurés sous forme textuelle pour la lecture. Rajoutons que l’alignement des schémas des réseaux viaires a été réalisé manuellement au moment du versement des réseaux dans la base de données.

Décalages planimétriques

En raison des erreurs de levé topographique et des erreurs résiduelles dues au géoréférencement des planches des atlas, les réseaux de rue présentent des décalages. L’amplitude de ces erreurs peut varier fortement selon les zones de l’espace parisien,

29. *Linestrings* du schéma simple feature access



(c) Extrait de la base de données spatiale et temporelle.

FIGURE 6.19 – Modélisation des réseaux viaires de Paris dans la base de données spatiale et temporelle

allant de quelques mètres dans le centre parisien de la rive droite à plus de 10 mètres pour les impasses, passages et cul-de sacs souvent présents dans les îlots anciens. Ces décalages peuvent être également dû aux élargissements des voies. En effet, puisque seuls les axes centraux des rues ont été vectorisés, un élargissement à droite ou à gauche d'une rue décale son axe. Deux situations sont présentées en figure 6.20, avec tous les réseaux superposés. À gauche, la partie nord de l'île Saint-Louis, presque inchangée, présente des décalages faibles. À l'inverse, la vignette de droite montre les décalages de rues et passages à l'intérieur d'un îlot, où les décalages planimétriques approchent 10 mètres. L'erreur est particulièrement importante pour les tronçons issus de l'atlas de Jacobet (en orange).

Identité des tronçons de rues

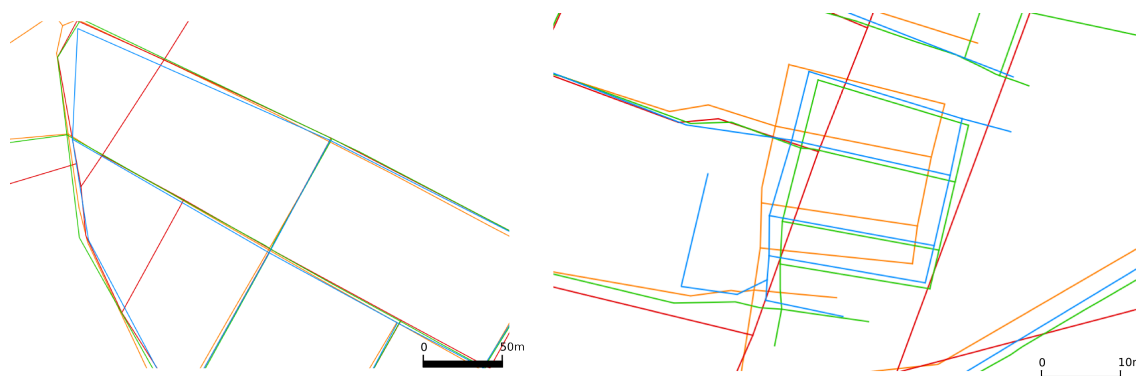


FIGURE 6.20 – Illustration des décalages planimétriques entre réseaux viaires sur l'île saint Louis et l'îlot de la trinité.

Nous utilisons dans cette section la géométrie et le nom de la rue comme éléments de l'identité d'un tronçon. Les attributs `'nom_entier'` et `'the_geom'` sont les deux membres de l'identité de tout tronçon de rue de Paris. L'attribut `'nom_entier'` contient le nommage complet de la rue, avec sa nature et sa dénomination (par exemple "rue Neuve Saint-Martin"). La géométrie du tronçon décrit la forme de son axe central et sa localisation.

Espaces parisiens étudiés

Pour présenter et discuter les résultats obtenus par notre approche sur les réseaux de rues de Paris, nous nous focalisons sur deux zones de Paris ayant subi des transformations différentes. De plus, nous souhaitons comparer les résultats avec une vérité terrain créée manuellement dont nous ne pouvons disposer que pour des espaces réduits en raison du volume de données. En effet, créer une vérité terrain consiste à créer dans la base de données spatio-temporelle un graphe de filiation pour les quatre plans. De plus, cette vérité terrain doit s'appuyer sur des connaissances historiques des transformations issues d'experts historiens.

Nous avons choisi de traiter deux zones de Paris de taille différente. Tout d'abord, nous nous intéresserons aux transformations de l'intérieur de l'îlot parisien de la trinité décrit dans (Gribaudi, 2009). Les différents états de l'îlot au travers de nos plans sont visibles en figure 6.21 présentant une transformation typique des anciens enclos religieux du centre de Paris. Nous traiterons plus particulièrement l'intérieur de cet îlot, dont les réseaux vectorisés dans la partie inférieure de la figure. Cet îlot est un espace occupé pendant l'Ancien Régime par l'enclos de l'Hôpital de la Trinité, supprimé en 1790. L'hôpital lui-même est détruit en 1817 et laisse place à des immeubles d'habitation et des ateliers d'artisans. La structure viaire de l'îlot change peu jusqu'aux interventions haussmanniennes au cours desquelles il est presque intégralement détruit en raison du percement du boulevard Sébastopol. Seule la partie Ouest de l'îlot échappe à la destruction, ce qui épargne des deux passages permettant d'atteindre ses rues internes. Le second espace que nous traiterons est le quartier des Arts et Métiers, entourant immédiatement l'actuel Conservatoire National des Arts et Métiers situé dans le 3^{ème} arrondissement. Les différents états du quartier sont illustrés en figure 6.22. Il s'agit ici aussi d'un ancien enclos religieux apparte-

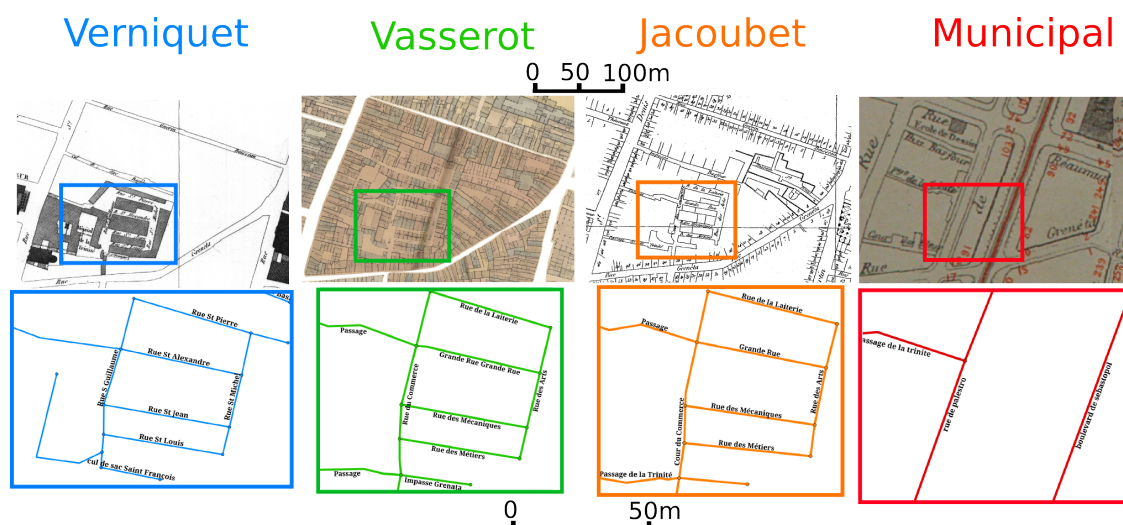


FIGURE 6.21 – Les états successifs de l’îlot de la trinité et les réseaux vectorisés correspondants.

nant à l’abbaye de Saint-Martin-des-Champs et de son abord immédiat. Le quartier présente cette fois des transformations plus variées. Tout d’abord, malgré l’ouverture de l’enclos vers l’extérieur avec la Révolution, les bâtiments de l’abbaye ont été pour la plupart conservés pour devenir le Conservatoire National. Une partie des rues est restée presque inchangée, en particulier dans la partie Nord du quartier. La partie Sud est à l’inverse presque entièrement détruite par le percement de la rue Réaumur entre 1854 et 1858. Seule la rue Aumaire, au Sud, est partiellement épargnée et existe encore aujourd’hui. La partie Est subit des transformations plus progressives, avec l’ouverture de rues et d’un marché à la place de l’ancien enclos dans un premier temps, puis des opérations d’alignement et de fusion des rues lors du Second Empire. À l’Ouest, la rue Saint-Martin est seulement alignée.

6.4.3 Fonctions d’agrégation des membres de l’identité des tronçons

Lors de l’évaluation de la vraisemblance d’un hyperarc durant l’étape de découverte des filiations, il est nécessaire d’agrèger les observations membres composant la queue et la tête de l’hyperarc. De cette façon, l’estimation de la vraisemblance d’un arc revient à un problème classique de recherche d’objets homologues. Cette vraisemblance étant calculée seulement à partir des attributs constituant l’identité des observations membres de l’hyperarc, il est seulement nécessaire d’agrèger ces attributs ainsi que le temps valide des observations. Dans le cas précis des réseaux viaires de Paris, tous les tronçons ont le temps valide de leur source, il n’y a donc pas besoin d’agrèger ces temporalités³⁰. Nous n’avons donc besoin de définir que deux fonctions d’agrégations, l’une pour les géométries des tronçons et l’autre pour les noms des rues. Nous présentons ici les deux méthodes choisies.

Fonction d’agrégation des géométries des tronçons de rue

30. Ceci est vrai car les hyperarcs ne peuvent être multi-sources.

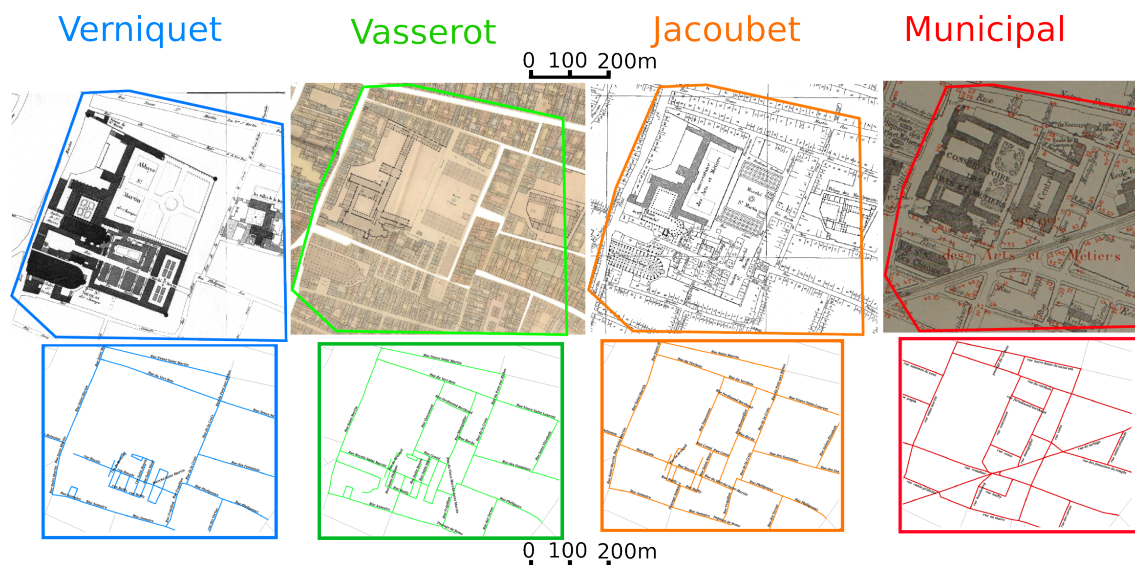


FIGURE 6.22 – Les états successifs du quartier des Arts et Métiers et les réseaux vectorisés correspondants.

Une première fonction d'agrégation doit permettre de fusionner les géométries des tronçons de rue constituant une extrémité d'un hyperarc de filiation. La façon dont nous parcourons l'espace de recherche des hypergraphes de filiation nous assure que chaque extrémité d'un arc est constituée de tronçons connectés spatialement. Cependant, la configuration de ces tronçons dans le réseau peut être quelconque. Il n'est donc pas toujours possible d'effectuer une simple union des géométries en une unique polyligne (c'est par exemple le cas dans une configuration en 'T').

Afin de pouvoir agréger toute configuration de tronçons, nous utilisons la méthode de création de *routes naturelles* introduit par Jiang *et al.* (2008). Une route naturelle est, dans un graphe routier, le chemin emprunté par un individu partant d'un point donné et parcourant le réseau en cherchant à conserver un trajet le plus rectiligne possible. Il s'agit donc d'agréger des tronçons consécutifs de façon à former un ensemble de lignes polygonales correspondant à des routes naturelles. L'algorithme proposé consiste à former ces chemins de manière itérative en agrégeant de proche en proche des tronçons connectés. Pour choisir à chaque intersection quels tronçons seront agrégés, l'auteur propose trois stratégies. La première, *best fit*, consiste choisir le tronçon à agréger aléatoirement parmi ceux formant l'intersection. La stratégie *self fit* consiste à conserver pour la route naturelle en cours de formation le tronçon formant un angle minimal. Enfin, la stratégie *every-best-fit* consiste à calculer le meilleur compromis à chaque intersection. Nous illustrons en figure 6.23b le résultat de la formation des routes naturelles sur un ensemble de tronçons en utilisant la stratégie *every-best-fit*. Cette méthode d'agrégation peut aboutir à la création de plusieurs polylignes agrégées lorsque la configuration des tronçons agrégés ne forme pas un chemin unique. Les mesures de similarité d'identité portant sur la géométrie doivent donc être capables de comparer des observations agrégées dont les géométries sont multiples.

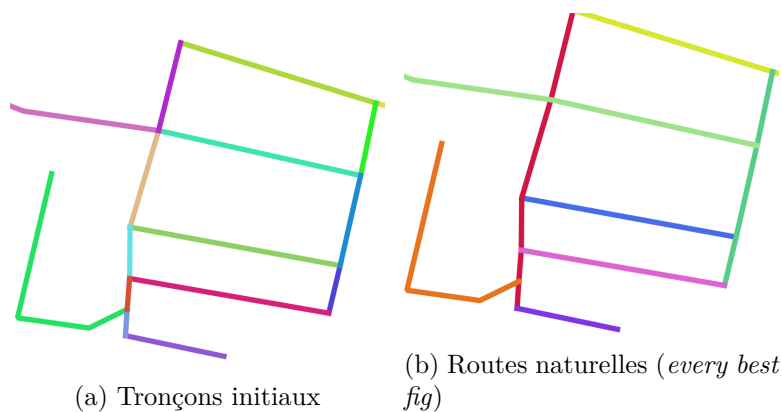


FIGURE 6.23 – Agrégation des tronçons de rues par création de routes naturelles. Chaque couleur correspond à une polyligne.

Fonction d'agrégation des noms de rues

Le second membre de l'identité des tronçons de rue est le nom de la rue à laquelle ils appartiennent. Lorsque les noms de rue des tronçons à agréger sont les mêmes, il suffit d'éliminer les doublons. Lors de la formation d'une observation agrégée, il est possible que des tronçons de rues de noms différents doivent être agrégés. Dans ce cas, la fonction d'agrégation des noms de rue renvoie un l'ensemble des noms de rues différents portés par les tronçons agrégés. Tout comme pour l'agrégation géométrique, les mesures de similarité portant sur ces noms doivent pouvoir comparer des ensembles de noms.

6.4.4 Mesures de similarité d'identité utilisées pour la découverte des relations de filiation

Nous avons défini le schéma des observations sur les tronçons de rue de Paris que nous utilisons dans cette section. Pour qu'il soit possible de découvrir les relations de filiations qui existent entre elles, nous devons définir des mesures de similarité d'identité permettant de comparer les attributs composant leur identité. Puisque nous avons défini l'identité d'un tronçon de rue comme étant formé du nom de rue et de sa géométrie, nous devons disposer de mesures de similarité pouvant comparer des géométries linéaires ainsi que des noms de rues représentés sous forme de chaînes de caractères.

Une fois ces mesures déterminées, elles doivent être intégrées dans des fonctions de masses de croyance respectant le modèle défini dans le paragraphe 6.2.3.

Concernant les mesures de succession temporelle, elles ne dépendent pas du cadre applicatif, ce qui nous permet de réutiliser les deux mesures définies pour le cadre général (voir le paragraphe 6.1.2).

6.4.4.1 Mesure de similarité entre géométries linéaires et fonction de masse associée

L'objectif est de définir des mesures capables d'estimer si les géométries des observations agrégées d'un hyperarc sont suffisamment similaires pour que l'on puisse

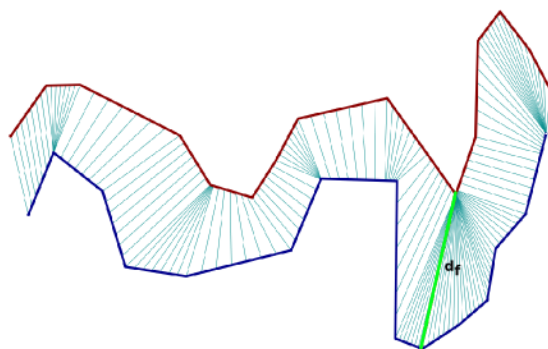


FIGURE 6.24 – Exemple de calcul de la distance de Fréchet entre deux lignes polygones.

considérer que les observations simples qui le compose sont affiliées. Au niveau des observations agrégées, cela revient à une mesure de similarité classique. Pour comparer deux lignes polygones, la plupart des méthodes d'appariement de données géographiques s'appuient sur la distance de Hausdorff ou de Fréchet (Alt et Godau, 1995). La distance de Hausdorff est une mesure du plus grand écart entre deux positions au sein des polygones. Pour expliquer le comportement de la distance de Fréchet entre deux courbes, l'analogie suivante est souvent proposée (Devogele, 2000; Mascret *et al.*, 2006) : un homme promène son chien en laisse, chacun suivant l'une des courbe sans pouvoir reculer mais en avançant ou en s'arrêtant chacun librement. La distance de Fréchet est alors la longueur minimale de la laisse qui permet au maître et au chien de parcourir leur courbe. La distance de Fréchet est définie dans un cadre continu, mais une discrétisation en a été proposée par (Alt et Godau, 1995) puis (Devogele, 2000). Devogele (2000) a par ailleurs montré que cette distance était mieux adaptée pour comparer des lignes sinueuses, et correspondant davantage à l'intuition humaine. La figure 6.24 illustre graphiquement le calcul de la distance de Fréchet d_f entre deux polygones. À partir de cette distance, nous pouvons construire une mesure de similarité d'identité entre géométries linéaire. Cette **mesure de Fréchet**, notée c_f est donnée entre deux observations o_1 et o_2 par

$$c_f(o_1, o_2) = e^{-\frac{K}{\epsilon_f}} \quad (6.14)$$

$$K = \sup\{d_f(U_g(o_1), U_g(o_2))\}$$

où ϵ_f est un paramètre de seuil permettant de fixer la distance à partir de laquelle la filiation entre o_1 et o_2 ne sera plus jugée vraisemblable d'après cette mesure. La limite entre vraisemblance et non vraisemblance étant fixée à 0.5 par la probabilité pignistique des hyperarcs décrivant des apparitions et des disparitions d'observations, ce paramètre peut être déterminé si l'on connaît la distance df_{max} à partir de laquelle une filiation n'est plus jugée vraisemblable. On a en effet $\epsilon_f = -df_{max}/\ln(0.5)$. La mesure de Fréchet peut donc être paramétrée par df_{max} dont le sens est plus intuitif. K est la plus grande distance de Fréchet calculée entre les lignes polygones composant les géométries des observations agrégées o_1 et o_2 .

La mesure de Fréchet étant une mesure de similarité d'identité, la fonction de masse qui lui est associée suit le modèle 1 de Appriou (1991). De plus, nous considérons cette mesure totalement fiable pour notre cas applicatif ($\alpha = 1$). Pour tout hyperarc

e d'un hypergraphe de filiation, cette fonction de masse est don :

$$\begin{cases} m_{e_i}(\omega_e) = c_f(o_1, o_2) \\ m_{e_i}(\neg\omega_e) = 1 - c_f(o_1, o_2) \\ m_{e_i}(\Omega) = 0 \end{cases}$$

où o_1 est l'observation agrégée de $Q(e_i)$ et o_2 celle de $T(e_i)$.

6.4.4.2 Mesure de similarité entre noms de rues

Pour mesurer une similarité entre noms de rues, nous réutilisons la similarité de chaînes de caractères proposée par Samal *et al.* (2004) fondée sur une distance de Damerau-Levenshtein. Cette dernière est une extension de la distance d'édition de Levenshtein prenant également en compte l'inversion de deux caractères adjacents. La similarité proposée, que l'on note s_r , consiste pour deux chaînes de caractère à découper chacune d'elle en mot, puis à calculer une matrice de similarité entre mots des deux chaînes à partir de la distance de Damerau-Levenshtein³¹. La similarité finale est la moyenne des valeurs des cellules maximisant les similarités inter-mots, sans réutiliser deux fois le même mot³².

La mesure de similarité entre noms de rues c_r est alors :

$$c_r(o_1, o_2) = e^{-\frac{K}{\epsilon_r}} \quad (6.15)$$

$$K = \sup\{s_r(U_{nom_entier}(o_1), U_{nom_entier}(o_2))\}$$

De la même manière que pour la mesure de Fréchet, ϵ_r est un paramètre de la mesure indiquant un seuil de vraisemblance. Il peut également être déterminé à l'aide d'une distance d'édition maximale dr_{max} .

Les noms de rues sont cruciaux pour savoir si une relation de filiation est une continuation ou de dérivation. Ils doivent impérativement être pris en compte au moment du typage des filiations. Lors de la phase de découverte, ces noms n'ajoutent toutefois que peu d'information à celle fournie par la mesure de Fréchet. Le fait que deux tronçons aient des noms différents ne fournit aucune information quant à leur liaison³³. Lorsque le nom est le même, cela permet seulement de supposer qu'ils sont membres de la même rue. La mesure de nom de rue est utile pour qualifier une relation, mais elle n'est pas suffisamment spécifique pour apporter une information significative par rapport à ce que peut exprimer la mesure géométrique. Il aura même tendance à dégrader localement l'efficacité du recuit simulé en augmentant les croyances en des relations concernant des tronçons d'une même rue mais éloignés. Pour cette raison, nous considérons que cette mesure n'est pas fiable pour la phase de découverte des relations ($\alpha = 0$). Toutefois, elle reste cruciale dans la phase d'étiquetage des relations découvertes. Dans les cas d'application présentés ici, l'appariement comporte donc seulement trois critères : deux temporels et un

31. La distance entre deux mots est transformée en une similarité par normalisation par le nombre de caractère du plus long mot.

32. Une seule cellule sélectionnée par ligne et par colonne

33. Même si la rue a changé de nom, des tronçons peuvent tout de même entretenir des relations de dérivation.

géométrique.

La fonction de masse résultante est alors :

$$\begin{cases} m_{e_i}(\omega_e) = 0 \\ m_{e_i}(\neg\omega_e) = 0 \\ m_{e_i}(\Omega) = 1 \end{cases}$$

La mesure n'étant pas fiable, la totalité de la croyance est placée dans l'ignorance. La mesure de noms de rues est un élément neutre du calcul de la vraisemblance d'un hypergraphe de filiation.

6.4.5 Fonctions de masse de croyance utilisées pour le typage des filiations

Nous définissons ici les fonctions de masse utilisées pour typer les relations de filiation issues de première étape du processus. Ces fonctions de masse peuvent s'exprimer sur le cadre de discernement $\Omega = \{\gamma, \delta, \epsilon\}$ défini pour chaque tronçon ou sur toute disjonction de 2^Ω . Nous avons choisi d'utiliser trois sources d'informations s'exprimant sur les deux attributs composant l'identité des tronçons :

1. une source a fondée sur une distance de Fréchet entre tronçon dont le but est de déterminer si la filiation est un cas d'erreur.
2. une source b comparant la longueur de deux tronçons de manière à distinguer une dérivation d'une continuation. L'hypothèse est la suivante : si deux tronçons affiliés sont de longueur très différentes, il s'agit d'une relation de dérivation.
3. une dernière source c 'information doit juger de la ressemblance des noms de rue des tronçons. Il y a continuation entre deux tronçons si la rue n'a pas changé de nom.

Les fonctions de masse associées à ces trois sources d'information sont données dans les paragraphes suivants.

Détection des relations erronées : a

Le rôle de la source d'information a est de détecter les relations de filiations erronées. Celles-ci sont filtrées par un critère de distance de Fréchet entre tronçons : au delà d'un seuil, la filiation est jugée erronée. Afin que les résultats soient cohérents avec la phase de découverte des filiations, cette distance doit égale à df_{max} . Pour deux tronçons o_1 et o_2 , la fonction de masse de croyance associée à la source a est :

$$\begin{cases} m_a(\{\delta, \gamma\}) = 1 \text{ si } d_f(o_1, o_2) \leq df_{max}, 0 \text{ sinon} \\ m_a(\epsilon) = 1 - m_a(\{\delta, \gamma\}) \end{cases}$$

Mesure des longueurs relatives des tronçons liés : b

Étant donné deux tronçons o_1 et o_2 de longueur respective l_1 et l_2 , la fonction de masse associée à la sources d'information b est :

$$\begin{cases} m_b(\gamma) = \alpha e^{\frac{-(1-\min(l_1,l_2)/\max(l_1,l_2))}{\epsilon_b}} \\ m_b(\delta) = 1 - m_a(\gamma) \\ m_b(\Omega) = 1 - \alpha \end{cases}$$

ϵ_b est un paramètre de seuil déterminé comme pour les autres fonctions de masse par $\epsilon_b = -d_b/\ln(0.5)$. Si un des tronçons est plus petit que l'autre d'un facteur $1/d_b$, la filiation sera considérée par la source d'information b comme une dérivation. α est un facteur de fiabilité fixé arbitrairement à 0.95. Abaisser légèrement la fiabilité de cette mesure permet de prendre en compte la croyance de la source d'information c même si les tronçons o_1 et o_2 sont identiques.

Source d'information sur les noms de rue : c

Le but de la source d'information c est de typer une relation de filiation entre deux tronçons o_1 et o_2 en s'appuyant sur leur nom de rue. La fonction de masse associée est :

$$\begin{cases} m_c(\gamma) = \alpha \text{ si } c_r(o_1, o_2) \leq dr_{max} \\ m_c(\{ \\ delta, \epsilon\}) = 1 - m_c(\gamma) \\ m_c(\Omega) = 1 - \alpha \end{cases}$$

6.4.6 Résultats sur l'îlot de la Trinité

Paramètres utilisés pour la phase de découverte des relations

L'approche que nous proposons repose sur un nombre important de paramètres empiriques. Nous donnons ici ceux utilisés pour la construction du graphe de filiation sur l'îlot de la trinité.

Un premier ensemble de paramètres concerne le recuit simulé. Les températures initiales ainsi que la solution initiales sont déterminées d'après la proposition de Ben-Ameur (2004). Le nombre d'itérations maximale a été fixé à 50000, avec un critère d'arrêt si aucune meilleure solution n'est trouvée avant 10000 itérations. Enfin, le paramètre de vitesse de décroissance de la température a été fixé arbitrairement à 0.9999 dans le but d'atteindre une température proche de 0 à 50000 itérations sans pour autant décroître trop rapidement.

Un second ensemble de paramètres concerne les critères de similarité d'identité et de succession temporelle pour la phase de découverte des filiations. Nous avons fixé empiriquement les paramètres suivants :

- seuil de la mesure de Fréchet : $df_{max} = 20$ mètres.
- seuil de la mesure de rang temporel : $drank_{max} = 1$. Au delà d'un saut au dessus d'une source géohistorique, une filiation sera considérée non vraisemblable.

Découverte des relations de filiation : visualisation des résultats d'appariement

Nous présentons ici le résultat de la phase de découverte des relations de filiation pour les tronçons de rues de l'îlot de la trinité pour les quatre sources cartographiques simultanément après 20000 itérations du recuit simulé. Une première figure (6.25) présente le résultat des appariements créés par l'algorithme. Pour des questions de lisibilité, le résultat est présenté source à source. La couleur des différents réseaux correspond au code couleur fixé en début de section. Les flèches en rose correspondent à au résultat de la transformation de l'hypergraphe de filiation créé en graphe simple mais non étiqueté. Il est ici géolocalisé pour permettre de visualiser les liens entre les observations géohistoriques dans l'espace géographique. Enfin, les numéros correspondent aux identifiants (gid) des tronçons.

Nous constatons tout d'abord que l'identification des relations de filiation est satisfaisante, aucun lien incohérent n'ayant été créé. Cependant, la solution contient plusieurs cas de liens $N : M$ qui apparaissent en raison des décalages planimétriques des données. Plus exactement, ce type de liens apparaît lorsque les tronçons des deux sources ne sont pas parfaitement face à face. Dans ces cas, l'agrégation des tronçons et la formation d'hyperarcs de cardinalité $N : M$ permet de diminuer la valeur de la mesure de Fréchet, ce qui augmente alors la vraisemblance de l'hyperarc.

La méthode a donc tendance à sur-apparier les réseaux en présence de décalages planimétriques faibles³⁴. Ce sur-appariement n'est pas nécessairement problématique car la phase de typage des relations permet d'éliminer au moins partiellement ce type de configuration.

Les tronçons non appariés sont également intéressants. Ainsi, les tronçons 2295 et 2494 du réseau de l'atlas de Verniquet ne sont liés à aucun autre tronçon. Le premier est situé à moins de 10 mètres du tronçon 1885 de l'atlas de Vasserot, et, d'après la mesure de Fréchet, sa liaison avec ce dernier est plus crédible qu'une disparition. Cependant, son non-appariement est dû au fait qu'il est préférable de lier le tronçon 2296 uniquement au tronçon 1885 de l'atlas de Vasserot. Notre méthode permet donc également de prendre en compte de façon implicite le voisinage des observations. Le tronçon 2494 n'est quand à lui pas apparié en raison de sa forme particulière. En effet, la distance de Fréchet entre ce tronçon et le tronçon 1877 est élevée : la croyance pour la distance de Fréchet de Verniquet à Vasserot est donc principalement placée sur la proposition de non-vraisemblance de la filiation. Enfin, notons que seuls deux tronçons de rues de l'atlas de Jacobet ont été appariés, la structure viaire ayant été totalement modifiée.

En réalité, le graphe présenté dans cette figure est la solution optimale du recuit simulé³⁵. La tendance au sur-appariement est intrinsèquement liée à la façon dont a été modélisée la mesure de Fréchet.

Pour éviter de tels sur-appariements, une possibilité est d'ajouter un coût aux hyperarcs de cardinalité $n : m$ en affaiblissant la croyance des mesures de similarité d'identité dans la proposition ω . De cette façon, de tels hyperarcs ne peuvent être créés que s'ils apportent augmentent significativement la vraisemblance de l'hypergraphe. Dans la figure 6.26, nous donnons en exemple le résultat obtenu pour la découverte des relations de filiations entre l'atlas de Verniquet et de Vasserot avec un coefficient d'affaiblissement de 0.90 pour les arcs de cardinalité $n : m$ pour la me-

34. Lorsque ceux-ci sont trop important, les observations ne seront pas appariées.

35. Sur 5 exécutions du recuit simulé, la solution optimale est systématiquement atteinte en moyenne à l'itération 6920. Pour 20 000 itérations, le temps de calcul moyen est de 12.8 secondes.

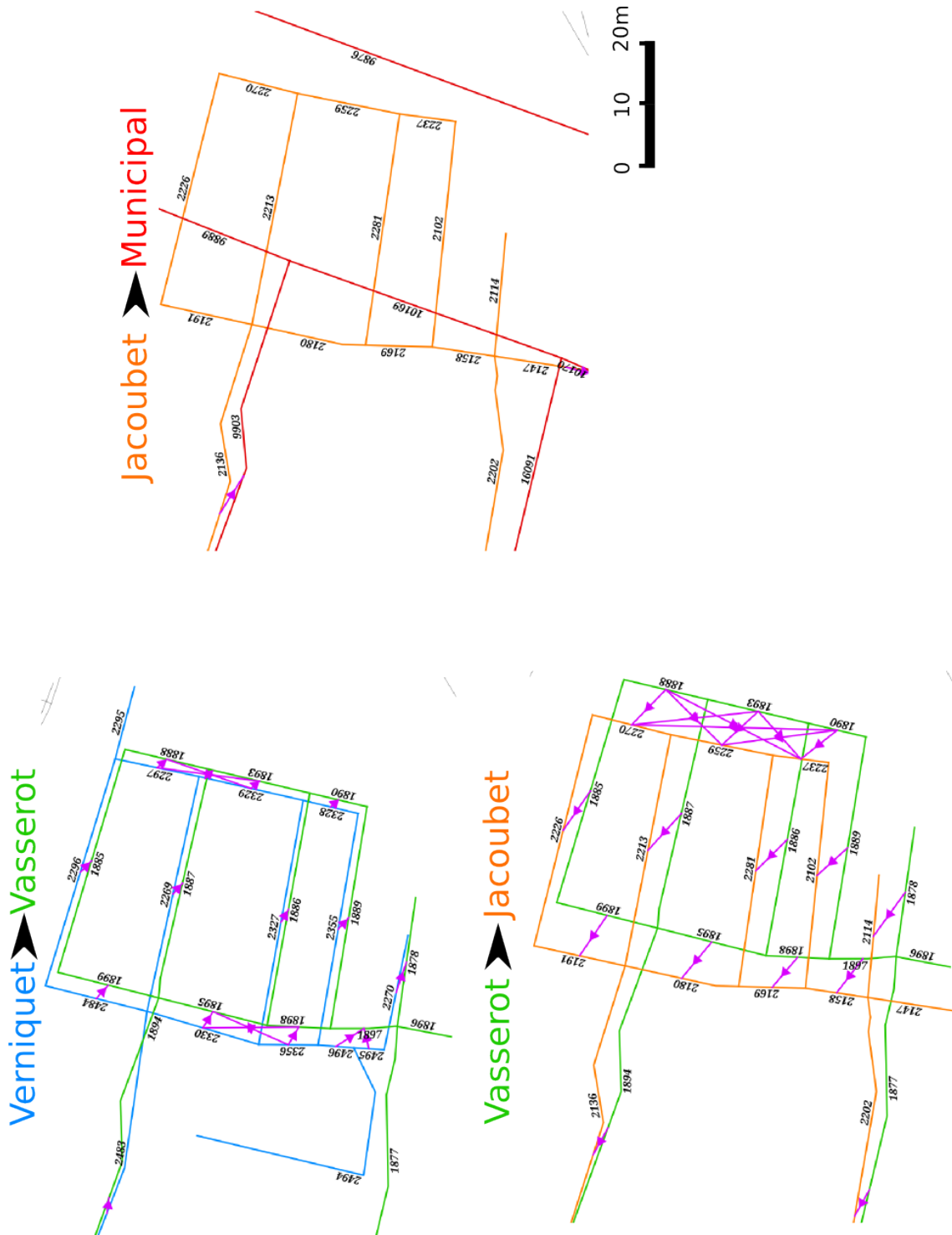


FIGURE 6.25 – Appariement résultant de la découverte des relations de filiations entre tronçons de l’îlot de la trinité.

sure de Fréchet³⁶. Les deux cas de sur-appariement présents dans la première version (à gauche) ne sont plus présents dans la solution optimale avec affaiblissement des arcs $n : m$. Ce paramètre est toutefois empirique et peut gêner la convergence de l'algorithme en empêchant la création d'arcs complexes dégradant momentanément la solution. De plus, nous avons dit que le sur-appariement n'était pas problématique car, à l'inverse du sous-appariement, il peut être corrigé par la phase de typage. Un comportement prudent consiste ici à ne pas pénaliser les hyperarcs de cardinalité $n : m$ afin d'améliorer la convergence du recuit simulé.

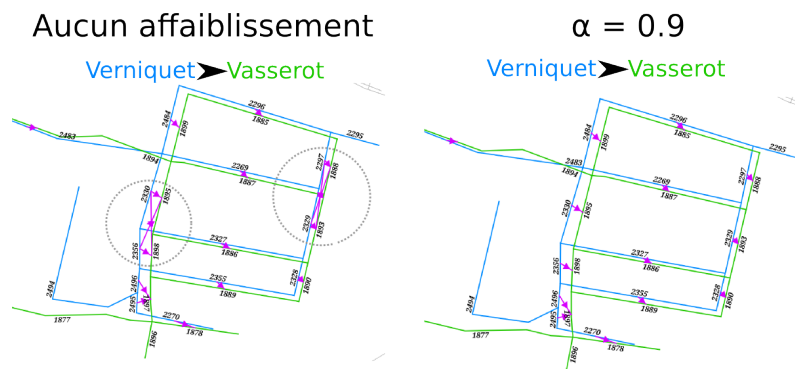


FIGURE 6.26 – Découverte des relations de filiation avec pénalité des hyperarcs de cardinalité $n : m$.

Découverte des relations de filiation : confrontation avec un algorithme d'appariement de réseaux classique

L'approche que nous avons proposé pour créer l'hypergraphe de filiation est une approche d'appariement de données géographiques. Il est alors intéressant de confronter les résultats que nous obtenons avec ceux d'une méthode d'appariement classique de réseaux. En effet, nous avons opté pour la généralité, mais une méthode spécifiquement dédiée à un type de données géographiques peut être plus efficace. Pour cela, nous avons créé une vérité-terrain pour l'îlot de la Trinité à partir des connaissances sur les transformations de ses rues issues de (Gribaudo, 2009) et (Lazare et Lazare, 1844). Nous avons ensuite choisi de comparer notre approche avec la méthode d'appariement de données géographiques proposée par Mustière et Devogele (2008). Celle-ci est en effet spécifiquement dédiée à l'appariement de réseaux de rues dont les niveaux de détail différent. Elle a de plus prouvé son efficacité (voir (Olteanu, 2008)) et est disponible dans la plateforme Geoxygene. Cette méthode d'appariement ne permet toutefois que des appariements entre deux sources de données. Nous l'avons donc exécutée pour chacun des couple de *snapshots* (Verniquet, Vasserot), (Vasserot, Jacoubet) et (Jacoubet, Municipal). Nous méthode est quand à elle appliquée à tous les *snapshots* simultanément. Pour comparer les deux méthodes, nous nous appuyons sur les mesures classiques de précision, rappel et f-mesure, définies par :

$$\text{precision} = \frac{\text{nombre de liens d'appariement corrects trouvés}}{\text{nombre de liens d'appariement trouvés}}$$

36. On a donc $m(\omega) = 0.90c_{\text{frechet}}$

TABLE 6.4 – Qualité de l'appariement sur l'îlot de la trinité

Méthode	Snapshots	Rappel	Précision	Fmesure
M&D	Verniquet → Vasserot	0.93	1	0.963
-	Vasserot → Jacoubet	1	1	1
-	Jacoubet → Poubelle	1	0.67	0.8
Notre méthode ($\alpha = 1$)	Verniquet → Vasserot	1	0.77	0.875
-	Vasserot → Jacoubet	1	0.63	0.77
-	Jacoubet → Poubelle	1	1	1
($\alpha = 0.9$)	Verniquet → Vasserot	1	1	1
-	Vasserot → Jacoubet	1	1	1
-	Jacoubet → Poubelle	1	1	1

$$\text{rappel} = \frac{\text{nombre de liens d'appariement corrects trouvés}}{\text{nombre de liens existants dans la vérité-terrain}}$$

$$\text{f-mesure} = 2 \times \frac{\text{precision} \times \text{rappel}}{\text{precision} + \text{rappel}}$$

La précision est une mesure du bruit présent dans le résultat. Le rappel indique quant à lui la pertinence du résultat : il est d'autant plus élevé que les liens de la vérité terrain ont été identifiés. Enfin, la f-mesure est une synthèse donnant une qualité globale de la solution.

Le tableau 6.4 présente le résultat de cette évaluation. Nous avons également donné le résultat des mesures pour la solution intégrant la pénalité sur les hyperarcs de cardinalité $n : m$. L'approche de Mustière et Devogele (2008) est noté "M&D".

Nous retrouvons dans les résultats la tendance de notre méthode au sur-appariement. En effet, la précision est relativement faible en raison des hyperarcs de cardinalité $n : m$ créés. Cet effet est entièrement corrigé lorsque l'on pénalise ce type d'hyperarcs. Le rappel est quant à lui systématiquement supérieur ou égal à celui obtenu avec la l'approche d'appariement de réseaux. Le nombre de liens d'appariement existants dans la vérité terrain est toutefois particulièrement faible (15 liens) tout comme le nombre de tronçons. L'utilisation d'une méthode automatique pour des cas de cette taille n'est pas pertinent. Il nous faut donc effectuer ces mesures sur un espace de plus grande taille.

Graphe de filiations non étiqueté et typage

Nous avons présenté jusqu'ici le résultat de la phase de découverte des relations de filiation selon l'angle de l'appariement de données géographiques. Le véritable objectif de cette phase est de produire un graphe G en décomposant le meilleur hypergraphe de filiation construit par le recuit simulé. Ce graphe, dont les arcs ne sont pas encore étiquetés, est présenté dans la figure 6.27. Notons que le graphe que nous présentons ici est celui obtenu en pénalisant les hyperarcs de cardinalité $n : m$ afin d'éliminer les cas de sur-appariement. Chaque tronçon est représenté par un cadre indiquant son identifiant et le nom de la rue à laquelle il appartient. Ce graphe étant une autre représentation du résultat illustré par les figures précédentes, nous retrouvons exactement les mêmes arcs. Nous pouvons déjà voir apparaître

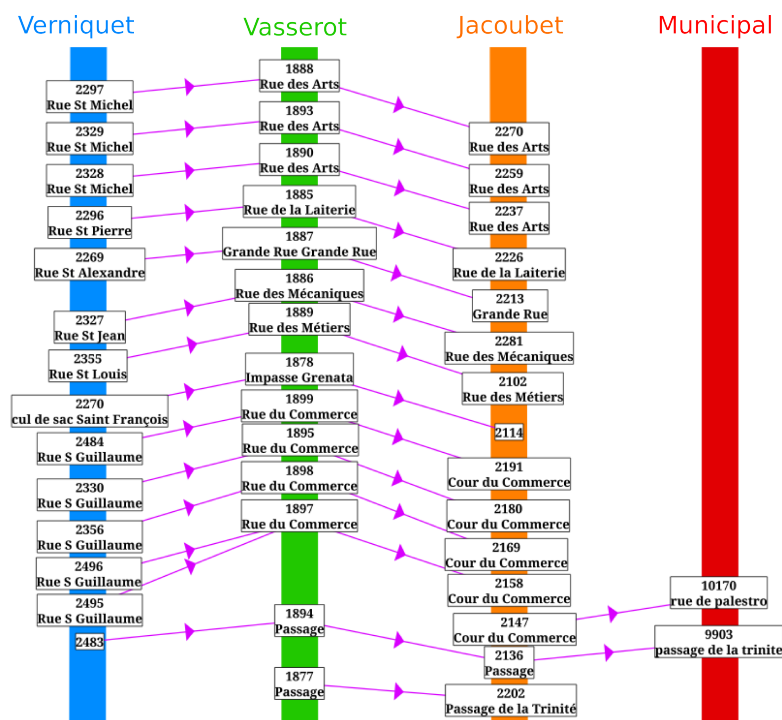


FIGURE 6.27 – Graphe G non étiqueté créé par la phase de découverte des relations de filiation pour l'îlot de la Trinité.

quelques phénomènes : la brisure haussmannienne qui détruit le quartier (lisible entre Jacoubet et l'atlas Municipal) ainsi que le changement presque intégral des noms des rues qui traduit le changement de statut de l'îlot après la Révolution. Nous appliquons ensuite le processus de typage des relations de filiation en utilisant les paramètres et mesures définies pour cet îlot. Ainsi, nous considérons qu'il y a dérivation entre deux tronçons si le nom de la rue change ou si la longueur des tronçons liés varie fortement, témoignant d'un découpage de la rue (en général par l'apparition d'une intersection). Enfin, les arcs sont considérés erronés s'ils sont situés à plus de 20 mètres les uns des autres. Le graphe de filiation G_f résultant de l'étape de typage est visible dans la figure 6.28. Les relations de dérivation sont marquées par un trait pointillé et les relations de continuation par un trait en plein. La probabilité pignistique du type choisi est indiquée par un dégradé de couleur et est de plus indiquée au dessus de chaque arc. Cette probabilité pignistique traduit certitude que nous avons sur le type de chaque lien. Le graphe formé fait apparaître les mutations des tronçons des rues de l'îlot. Ainsi, la totalité des tronçons de l'atlas de Verniquet sont en relation de dérivation avec ceux de l'atlas de Vasserot puisque le nom des rues changent. La probabilité pignistique de ces dérivation est toutefois peu élevée car les tronçons sont géométriquement très similaires. Entre les atlas de Vasserot et de Jacoubet, l'état de l'îlot ne varie plus. Les seules relations de dérivation identifiées sont dues à des données manquantes. Enfin, tous les tronçons disparaissent à l'exception du tronçon 2136 et 2147. Le premier existe toujours dans l'atlas municipal, mais une dérivation a été identifiée en raison du nom partiellement manquant dans l'atlas de Jacoubet. Concernant le tronçon 2147, la dérivation est

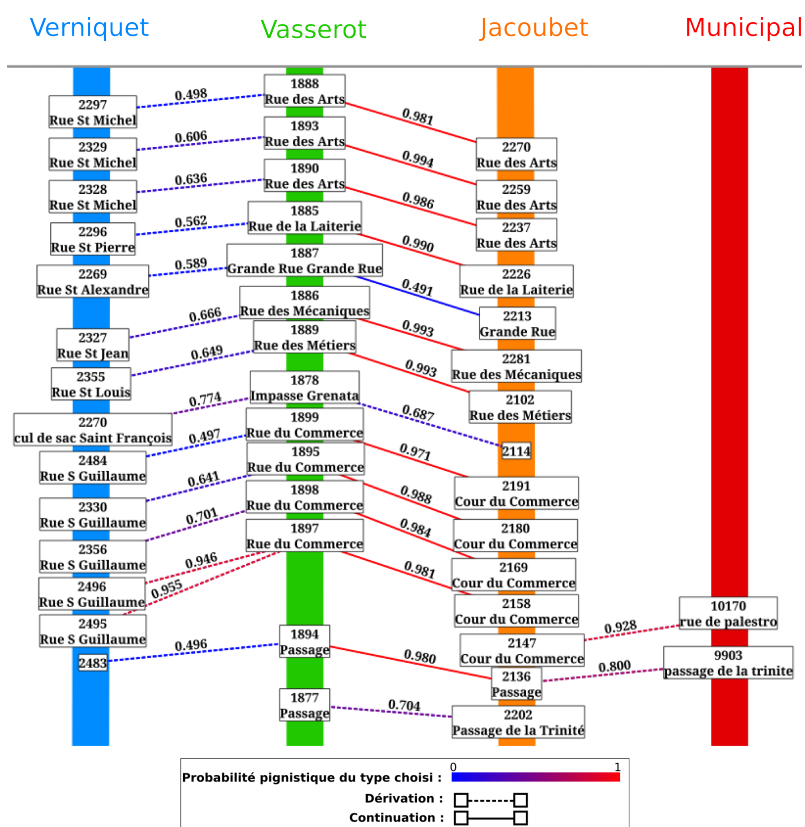


FIGURE 6.28 – Graphe de filiation G_f après typage des relations de filiations découvertes pour l'îlot de la trinité.

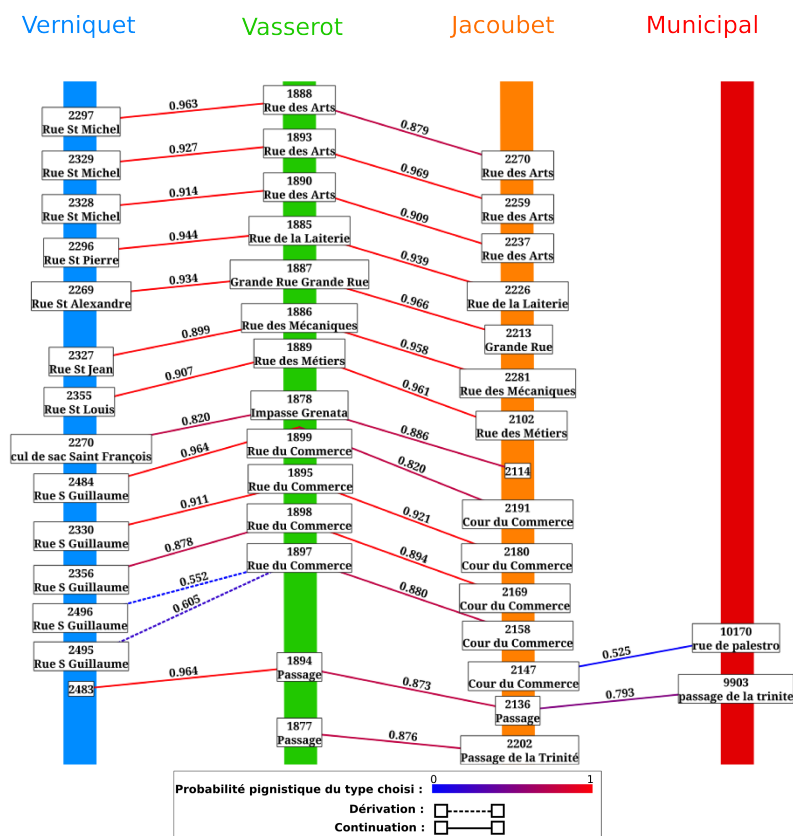


FIGURE 6.29 – Graphe de filiation G_f après typage des relations de filiations découvertes pour l’îlot de la trinité, sans considérer le nom des rues.

due au fait que la rue de Palestro occupe l’espace de la Cour du Commerce. Pour illustrer les différents graphes de filiation qui peuvent être créés à partir d’un même hypergraphe de filiation, nous avons ré-appliqué le processus de typage en considérant cette fois seulement la morphologie du réseau. Pour que le nom des rues n’intervienne plus dans le choix du type, nous avons modifié la mesure associée afin qu’elle place toute sa croyance systématiquement dans l’ignorance totale. Le résultat est visible dans la figure 6.29. Cette fois, la majorité des relations sont des continuations, la géométrie des tronçons variant peu jusqu’à leur destruction (sauf décalages planimétriques). Seuls les tronçons de la rue Saint-Guillaume sont en relation de dérivation car la disparition du tronçon 2494 (visible en figure 6.25) provoque leur fusion. Notons que le nom de la rue n’intervenant plus dans le typage, la relation entre la Cour du Commerce et la rue de Palestro est cette fois identifiée comme une continuation.

6.4.7 Résultats sur le quartier des Arts et Métiers

Bien que les résultats sur l’îlot de Greneta sont satisfaisant, il s’agit d’un petit espace composé de peu de tronçons. Nous appliquons ici notre approche à un espace de plus grande taille (251 tronçons répartis dans 4 *snapshots*) dans le but d’évaluer sa qualité dans des configurations plus complexes. Les résultats présentés ici ont été obtenus après 100000 itérations du recuit simulé, avec une température initiale

$T_0 = 1.4$ et un facteur de décroissance de température $\mu = 0.9999$. Le reste des paramètres est identique à l'îlot de la trinité.

Visualisation des résultats d'appariement

Le résultat de la phase de découverte des filiations est présenté sous la forme d'un résultat d'appariement en figure 6.30. Les relations de filiations sont indiquées par des flèches en rose. Afin de discuter les résultats, nous allons effectuer quelques zooms sur certaines zones pour lesquelles l'algorithme de découverte des relations ne permet pas de trouver de bonne solution. La figure 6.31 pointe trois situations entre les atlas de Jacoubet et Municipal dans lesquelles l'algorithme n'a pas permis de détecter les relations de filiation. Le détail de ces trois zones et la vérité terrain correspondante est visible en figure 6.32. La décomposition des hyperarcs de la vérité terrain sont tracés en noir, ceux de la solution trouvée en rose. Les cas *a* et *c* correspondent à des processus spatio-temporels de réallocation. Dans le premier cas, la rue (il s'agit de la rue Saint-Martin) reste inchangée entre les deux atlas, mais plusieurs nouvelles rues apparaissant avec les percées haussmanniennes viennent créer de nouvelles intersections et ainsi diviser le tronçon 1202 de l'atlas de Jacoubet en 3 tronçons. Dans le même temps, le tronçon 1191 de Jacoubet disparaît. Dans le cas *c*, la situation est du même type : les tronçons 3665, 3667, 148, 137 et 138 de Jacoubet, formant initialement deux rues (rue de Frépillon et rue de la Croix) sont alignées et rassemblées en une rue unique (rue Volta) composée des tronçons 10199, 9858 et 9860. Enfin, le cas *b* correspond au remplacement d'une ancienne rue (rue Royale) par un tronçon de l'actuelle rue Réaumur. L'échec de l'algorithme de découverte des filiations est ici dû à un problème de convergence. Dans les trois cas, les tronçons à appairer sont de taille très différentes, avec de petits tronçons devant être affiliés à des tronçons de grande taille (par exemple les tronçons 143 et 9871 du cas *b*). Les hypergraphes étant construits pas à pas pendant le parcours de l'espace de recherche en commençant par la création d'un hyperarc de cardinalité 1 : 1, les hyperarcs des cas *a*, *b* et *c* ne peuvent être atteints que par la complexification d'un hyperarc créé initialement entre deux tronçons. Puisque ceux-ci sont très différents, la mesure de Fréchet va considérer la filiation de ces deux tronçons comme non vraisemblable et il sera alors momentanément préférable de ne pas les affilier. Seule l'extension de l'hyperarc initial vers un arc de cardinalité $n : m$ permettra de trouver la bonne solution. L'absence de filiation est donc ici un cas de minimum local qui devrait être surmonté par l'algorithme. Or, celui-ci ne parvient pas à sortir de ce minimum local pour deux raisons :

- plus les tronçons initialement liés sont éloignés et de taille différente, plus la dégradation de la fonction initiale sera importante. Plus encore, une dégradation -de moins en moins importante toutefois- aura lieu pour toutes les étapes intermédiaires menant à l'hyperarc $n : m$. Cet hyperarc ne peut donc être trouvé que lorsque la température du recuit est élevée,
- lorsque la température est élevée, l'algorithme effectue une marche aléatoire dans l'espace de recherche. Il devient donc improbable de créer l'hyperarc souhaité³⁷. Même si cet hyperarc est trouvé, le parcours de l'espace de recherche à température élevée peut aboutir à sa destruction.

37. La probabilité est d'autant plus faible que la taille du problème augmente.



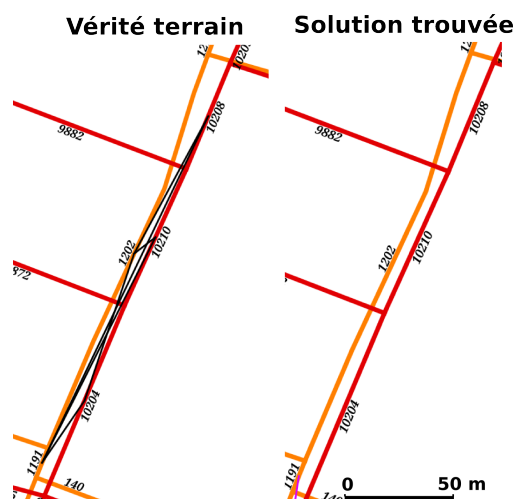
FIGURE 6.30 – Appariement résultant de la découverte des relations de filiations entre tronçons pour le quartier des Arts et Métiers.



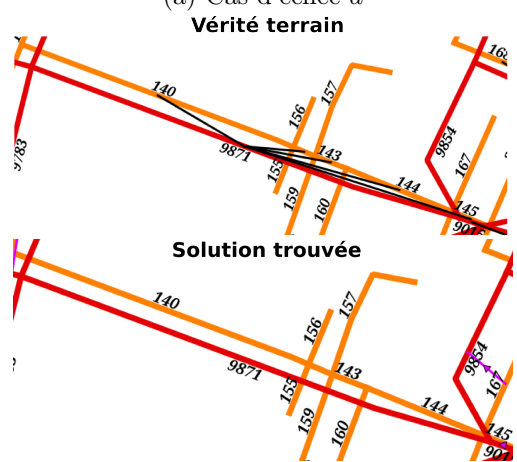
FIGURE 6.31 – Cas d'échecs de l'algorithme de découverte des relations de filiation.

Ces deux effets se cumulent donc et gênent fortement la convergence de l'algorithme sur les cas de processus spatio-temporels entre tronçons de géométries très différentes.

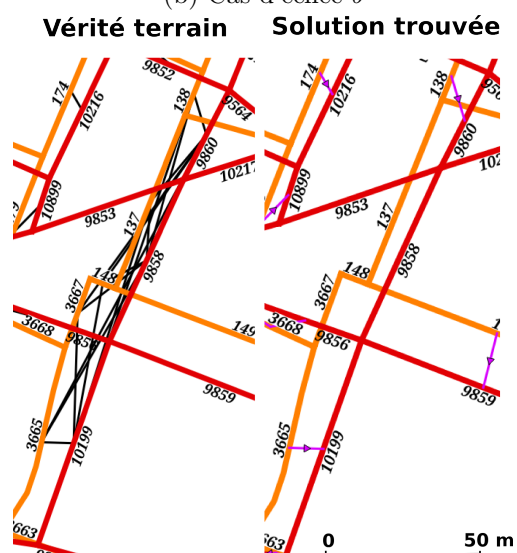
Un second type de cas d'échec est visible en figure 6.33, lié cette fois aux décalages planimétriques des réseaux viaires. Nous nous situons cette fois dans la partie centrale de l'ancien enclos de Saint-Martin des Champs. Le réseau issu de l'atlas de Jacoubet, figuré en orange, est décalé en moyenne de 10 mètres sur l'ensemble de cet espace, créant un chevauchement avec le réseau issu de l'atlas de Vasserot (en vert). Le décalage est tel que les tronçons les plus proches dans les deux atlas ne sont pas ceux affiliés dans la vérité terrain. La localisation des tronçons de l'atlas de Jacoubet corrigée manuellement, visible dans l'encart de la même figure, permet de se rendre compte de l'effet de chevauchement des deux réseaux. Dans notre approche de découverte des filiations, nous n'avons pas pris en compte les décalages planimétriques entre observations issues de différents *snapshots*. La filiation étant fondée sur des distances, les tronçons les plus proches sont choisis, ce qui peut être faux lorsqu'il y a des décalages importants. De plus, comme nous ne prenons pas en compte le nom de la rue à laquelle appartient le tronçon, seule la géométrie apporte une information sur l'existence d'une relation de filiation entre deux tronçons. Les erreurs visibles sur la figure 6.33 sont dues à cet effet. Par exemple, le tronçon 3220 encadré en vert devrait être affilié au tronçon 155 (encadré en orange), et non au tronçon 159 pourtant plus proche en raison du décalage du réseau. Afin de corriger ce type d'erreur, une possibilité serait d'intégrer au sein même de la découverte des filiations une étape de recalage planimétrique.



(a) Cas d'échec a



(b) Cas d'échec b



(c) Cas d'échec c

FIGURE 6.32 – Détail des situations d'échec de la phase de découverte des relations de filiation pour le quartier des Arts et Métiers.

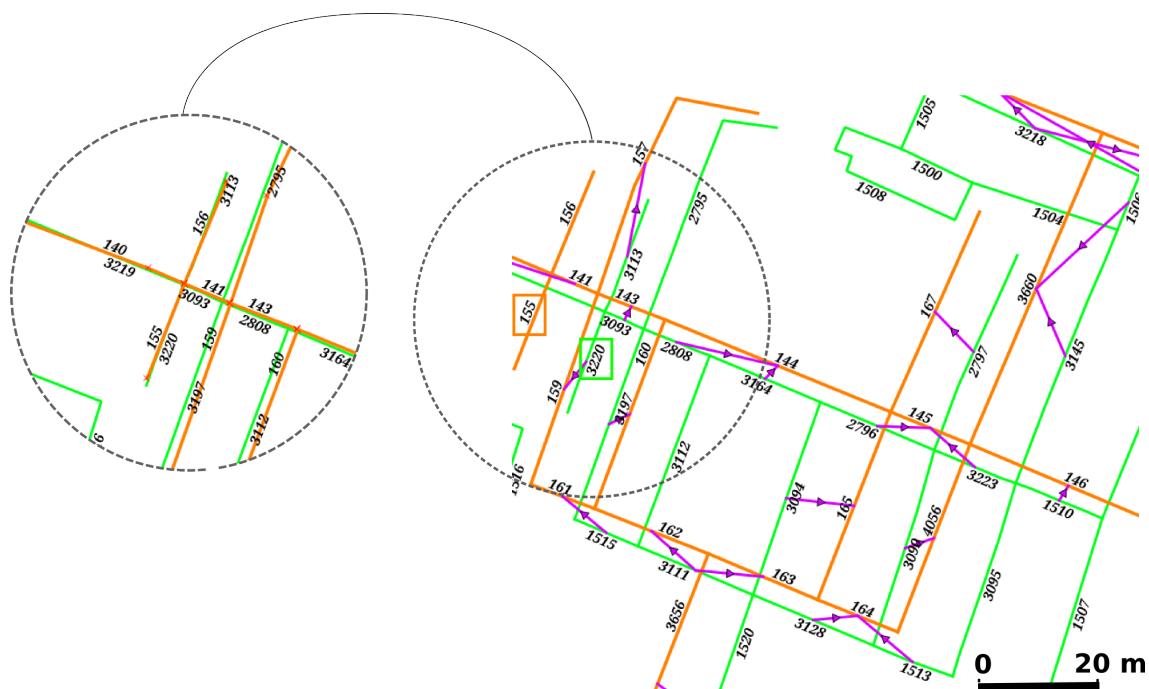


FIGURE 6.33 – Faux positifs lors de la découverte des filiations, dus au décalage planimétrique des données. À gauche, les décalages planimétriques de l’atlas de Jacoubet sont réduits manuellement par rapport à celui de Vasserot.

Comparaison avec la méthode d’appariement de Mustière et Devogele (2008)

De la même façon que pour l’îlot de la trinité, nous avons comparé les résultats obtenus pour le quartier des Arts et Métiers avec la méthode d’appariement de réseaux de Mustière et Devogele (2008). Les mesures sont présentées dans le tableau 6.5. Tout comme pour l’îlot de la trinité, nous obtenons une valeur de rappel systématiquement meilleur que l’approche spécialisée. Le rappel décroît cependant fortement entre les atlas de Jacoubet et Municipal en raison des problèmes présentés dans les paragraphes précédents. Il en résulte un sous-appariement qui empêchera la création d’un graphe de filiation satisfaisant et demande une correction manuelle. L’approche de Mustière et Devogele est peu efficace pour détecter les relations de filiation entre ces deux derniers atlas car elle n’est pas adaptée à des réseaux dont la structure change. En effet, l’objectif de cette approche étant d’apparier des réseaux à différents niveaux de détail mais représentant les mêmes entités du monde réel, elle suppose que les topologies des réseaux appariés sont proches et utilise cette topologie pour effectuer l’appariement. Or, nous sommes dans un cas différent où les réseaux ne représentent pas nécessairement les mêmes entités : la topologie des réseaux n’est donc pas toujours d’un grand secours et peut même nuire à l’appariement. Concernant la précision, notre approche a toujours tendance à sur-apparier les réseaux en créant des hypergraphes complexes.

TABLE 6.5 – Qualité de l'appariement sur le quartier Saint Martin

Méthode	Snapshots	Rappel	Précision	Fmesure
M&D	Verniquet → Vasserot	0.836	0.98	0.90
-	Vasserot → Jacoubet	0.75	0.86	0.8
-	Jacoubet → Municipal	0.30	0.90	0.45
Notre méthode ($\alpha = 1$)	Verniquet → Vasserot	0.95	0.90	0.93
-	Vasserot → Jacoubet	0,76	0.71	0.736
-	Jacoubet → Municipal	0.52	0.74	0.61

6.5 Conclusion

Dans ce dernier chapitre nous avons présenté une approche générique permettant de construire des graphes de filiations à partir d'observations géohistoriques structurées au sein d'une base de données spatio-temporelle. Cette approche s'appuie sur un processus en deux temps. Tout d'abord, les observations géohistoriques sont appariées afin de détecter les observations dont les identités sont liées. Nous proposons pour cela une nouvelle méthode d'appariement de données géographiques par optimisation discrète entre un nombre quelconque de sources de données. Cette méthode s'appuie sur la théorie des croyances introduite par Dempster (1968) et utilisée dans le cadre de l'appariement de données géographiques par Olteanu (2008). Dans un second temps, un graphe de filiation est créé à partir du résultat de l'appariement à l'aide de connaissances imparfaites sur la nature des relations de filiation entre observations. Ceci nous permet de générer des graphes de filiation automatiquement à partir d'une définition de l'identité d'une observation fournie par l'utilisateur.

Nous avons développé un prototype expérimental nous permettant de tester l'approche sur les réseaux de rues de Paris vectorisés et présentés dans le chapitre 3 puis stockés dans une base de données spatiale et temporelle dont le schéma a été explicité au cours du chapitre 4. Les premières expérimentations de notre approche démontre sa capacité à identifier des relations de filiation entre observations géohistoriques vectorielles issues d'un nombre quelconque de sources. Les résultats sont ainsi encourageants sur les données de réseau de rues utilisés. La méthode recherchant des agencements d'objets géographiques décrivant des transformations (ou des persistences), elle se montre plus adaptée qu'une approche d'appariement spécialisée dédiée à l'identification d'objets homologues choisie dans la littérature, tout en traitant simultanément plusieurs sources de données. Notre approche est également fortement générique, lui permettant d'être appliquée à des observations de nature différentes. L'étape d'étiquetage permet quant à elle une maîtrise relativement fine de la notion d'identité d'une observation géohistorique. Elle offre en particulier la possibilité de créer plusieurs graphes géohistoriques à partir d'un même résultat d'appariement correspondants à une conception différente de l'identité des observations en jeu.

Les expérimentations mettent également en évidence certaines limites. Tout d'abord, le processus proposé est relativement complexe et s'appuie sur plusieurs de paramètres empiriques, ce qui questionne réutilisabilité, ou, tout du moins, à sa robustesse. De plus, nous avons pointé des failles dans la convergence de la méthode qui gênent le passage à l'échelle. Ces limites ouvrent plusieurs pistes d'amélioration de l'approche de construction de graphes géohistoriques.

La première concerne le passage à l'échelle de l'approche. Les cas d'application sont ciblés sur des zones restreintes de Paris, en raison de la complexité de la représentation visuelle des résultats mais également de difficultés de convergences du recuit simulé³⁸. Le parcours du voisinage d'une solution au cours du processus d'optimisation est un point critique conditionnant à la fois la rapidité de l'optimisation et la qualité du résultat. À une itération donnée, le nombre de modifications applicables au graphe en cours de construction est élevé mais seules quelques modifications sont

38. Ces difficultés ont été abordées au cours des expérimentations.

réellement pertinentes : seule une petite portion de l'espace de recherche est réellement intéressante. Actuellement, aucune notion de pertinence n'est incluse dans l'étape de génération et de choix d'une nouvelle solution, la nouvelle solution étant tirée selon une loi uniforme. Le risque est alors de passer un temps important dans des zones non pertinentes de l'espace de recherche. Pour pallier ce problème, plusieurs pistes sont envisageables. Une façon de guider le parcours de cet espace serait de remplacer la loi uniforme par une loi ad-hoc privilégiant les solutions décrivant des processus spatio-temporels dont la structure est crédible. Cette crédibilité dépend bien sûr des données et peut varier selon le lieu et l'époque des entités du monde réel considérées. La division d'une parcelle en quelques parcelles plus petites est un processus crédible, tandis que l'existence d'un large processus $N : M$ mêlant plusieurs dizaines de parcelles est relativement peu probable. Pourtant, ces deux solutions ont actuellement la même chance d'être atteintes. Une loi de probabilité adaptée à un espace, un temps et un type d'entités du monde réel doit être construite à partir de connaissances d'experts, ce qui pourrait être obtenu par une étape d'apprentissage supervisé permettant d'entraîner la méthode pour un usage donné.

Le recuit simulé étant une méthode d'optimisation locale et le problème d'appariement étant en réalité composé d'une multitude de sous-problèmes d'identification de processus spatio-temporels, l'application de l'approche à des espaces de plus grande taille (Paris, île de France, France entière, etc.) pose clairement des questions sur le temps de calcul nécessaire à l'obtention d'une solution de bonne qualité. Il est utile de s'interroger sur la possibilité de parallélisation de l'approche pour la construction de graphes de grande taille, lorsqu'il s'agit d'apparier un nombre important d'observations. Ce problème se pose tant dans la largeur (traiter des portions d'espace géographique de grande taille) que dans la profondeur (lorsque de nombreuses sources sont disponibles sur une portion d'espace donnée). Indirectement, la question de la parallélisation de l'approche pose la question du découpage du graphe à construire en sous-problèmes concernant une étendue spatiale et temporelle et la gestion des frontières entre ces sous-problèmes. Il s'agit ici d'une question ouverte qui nécessitera un travail ultérieur.

Une seconde perspective concerne l'amélioration de la qualité des solutions proposées. Nous avons vu que, bien que l'approche soit robuste aux imperfections de faible ampleur (décalages planimétriques notamment), la qualité des résultats baisse fortement lorsque celles-ci sont importantes. Concernant le cas précis des décalages planimétriques, il serait particulièrement intéressant d'inclure pendant l'optimisation des phases de recalage s'appuyant sur les relations de filiation déjà découvertes. Ainsi, il s'agirait de procéder à la fois à la création d'un graphe géohistorique et au recalage géométrique des données. Ce type de procédé a d'ailleurs déjà été appliqué pour l'appariement de deux réseaux par Li et Goodchild (2011b). Ceci nécessite toutefois de déterminer le moment adéquat de l'application du recalage pendant l'optimisation. On peut par exemple imaginer d'attendre que la qualité des solutions ne stagne ou bien de l'appliquer à itérations fixes. Il faut également déterminer s'il est alors préférable de poursuivre la descente de température ou bien de commencer une nouvelle phase de descente rapide à haute température³⁹. Le choix d'une stratégie doit être guidé par des expérimentations.

39. en utilisant par exemple une stratégie de *reannealing*. Pour cela, voir notamment (Ingber, 1989)

L'extension de l'approche à des observations de nature diverse est également une perspective intéressante. En effet, nous avons pour l'instant considéré seulement le cas où les schémas de toutes les observations traitées sont alignés et que celles-ci sont de même nature. Or, des processus spatio-temporels peuvent tout à fait concerner des observations de nature différentes. Il est par exemple courant, durant la première moitié du XIX^e, que des propriétaires parisiens fassent ouvrir de nouvelles rues au sein d'îlots bâtis, remplaçant une partie de l'espace bâti par le tracé d'une rue. Se trouvent donc mêlés bâtiments, parcelles et voies dans un processus global de percement de rues. Or, celui-ci n'est représentable au sein d'un graphe géohistorique qu'à condition que celui-ci comporte des observations de différentes natures. Ceci implique d'étendre l'approche pour considérer des processus spatio-temporels plus complexes porteurs d'une sémantique particulière (le percement d'une rue par exemple). Cela oblige également à étendre la définition des mesures pour qu'elles soient capables de comparer les identités d'observations de natures différentes.

Une dernière perspective porte sur la détermination de l'identité des observations et, par là, des paramètres et fonctions de masses de croyances nécessaires à l'étape d'étiquetage. En effet, cette définition est difficile car elle nécessite une connaissance précise de l'identité des observations, notion pourtant régulièrement floue qu'il n'est pas possible de décomposer en attributs comme nous le faisons ici. Pour aider à cette définition, une première possibilité consiste à procéder par apprentissage supervisé des paramètres des fonctions de masse de croyance à partir d'exemples de relations de continuation et de dérivation soumis à des experts historiens. Même appris, il est tout à fait possible que ces paramètres ne soient valides que pour une portion d'espace donnée ou seulement pour un usage précis.

Conclusion et perspectives

Rappel des objectifs

Dans ce travail de thèse, nous nous intéressons à la création de bases de données spatio-temporelles sur l'espace urbain ancien à partir de plans et cartes topographiques pouvant être utilisés dans le cadre de recherches en histoire urbaine. L'objectif principal de cette recherche est l'intégration des objets géographiques représentés par ces sources cartographiques au sein de graphes spatio-temporels permettant d'analyser les transformations des structures d'une ville. Les sources cartographiques anciennes, ou *sources géohistoriques*, ont la particularité d'être fortement hétérogènes et imparfaites. De plus, leur fiabilité doit constamment être remise en question pour qu'elles soient utilisables au sein d'études historiques. Ces différentes imperfections ont des effets sur les *informations géohistoriques* représentées par ces sources, augmentant notamment les incertitudes sur les formes urbaines et leurs évolutions.

Étant donné ces éléments, deux sous-objectifs sont visés plus particulièrement dans cette thèse :

1. Une base de données spatio-temporelle dédiée à l'histoire n'est utilisable qu'à condition que les données qu'elle comporte soient bien connues et que leurs imperfections soient identifiées et modélisées.

Le premier objectif consiste à identifier et modéliser les imperfections des sources géohistoriques et des objets géographiques qu'elles représentent.

2. Les sources cartographiques anciennes présentent un état d'une ville pendant une période de temps plus ou moins bien connue. Elles forment des couches spatio-temporelles, ou *snapshots*. Former des données spatio-temporelles à partir des objets géographiques qu'elles représentent implique d'identifier les transformations ou la persistance, au fil des couches, des entités du monde réel correspondant à ces objets.

Le second objectif est donc de proposer une approche permettant de lier les objets géographiques représentés dans les sources géohistoriques afin de former des objets spatio-temporels décrivant les transformations de l'espace urbain ancien.

Afin de proposer des réponses adaptées à des données réelles et dans le but de valider les solutions proposées, le cas applicatif des transformations de la ville de Paris du XVIII^e au XIX^e siècle sert de fil d'Ariane à l'ensemble du document. Pour cela, nous

traitons 5 grands plans topographiques de Paris ayant eu des rôles centraux dans l'histoire de la cartographie parisienne.

Résumé des propositions

Afin d'atteindre les deux objectifs identifiés, nous présentons un ensemble de propositions que l'on peut organiser ainsi :

- **Identifier les imperfections spatiales et temporelles des sources géohistoriques.** Nous proposons tout d'abord une approche d'analyse fine des sources cartographiques anciennes s'appuyant à la fois sur une étude qualitative et historique des sources et sur une étude quantitative des distorsions géométriques des cartes et des objets géographiques qu'elles renferment. L'analyse qualitative permet de déterminer la période pendant laquelle les plans peuvent être supposés de façon fiable l'espace de la ville⁴⁰, que l'on identifie comme la période de levé topographique. Elle permet également de mettre en évidence des liens entre plans, tels que réutilisations, reprises, imitations, qui apportent des connaissances sur la fiabilité du contenu de ces sources. L'analyse quantitative s'appuie, quant à elle, sur le géoréférencement des différents plans, dans la lignée des travaux de Jenny et Hurni (2011). Nous montrons ainsi comment le géoréférencement des plans de Paris peut être utilisé pour analyser les distorsions géométriques des différents thèmes cartographiques d'un plan. Le couplage de ces deux analyses permet alors de connaître et de mesurer les incertitudes temporelles et les imperfections spatiales d'une source géohistorique.
- **Modéliser des informations géohistoriques dans un SIG :** Après extraction par vectorisation manuelle⁴¹ des réseaux viaires des rues de Paris sur les plans topographiques géoréférencés afin de constituer un ensemble de données géohistoriques, nous proposons un modèle relationnel de base de données spatiale et temporelle visant à stocker ces données tout en conservant les imperfections identifiées dans l'étape précédente. Ce modèle organise les données vectorisées, nommées *observations géohistoriques*, en couches temporelles correspondant à une source géohistorique. Il s'appuie en outre sur les standards de l'information géographique ISO 19108, ISO 19109 et ISO 19107 qu'il étend pour leur intégrer les spécificités des données issues de sources géohistoriques. En particulier, nous proposons une modélisation des temporalités des sources et des observations fondés sur des sous-ensembles flous permettant de représenter l'incertitude sur la période durant laquelle ces observations représentent une entité du monde réel existante.
- **Modéliser et identifier les transformations de l'espace urbain :** Nous proposons enfin un ensemble de solutions permettant de construire des données spatio-temporelles à partir des observations géohistoriques stockées dans la base de données spatiale et temporelle. Pour cela, nous introduisons tout d'abord un modèle de graphe spatio-temporel étendant un modèle de la lit-

40. En réalité, le temps valide des plans est difficile à déterminer avec précision. On se contente de l'approximer par la période de levé topographique.

41. Nous avons fait le choix d'effectuer cette vectorisation de façon manuelle, mais on pourrait tout à fait envisager de reproduire ces travaux sur la base de données acquises de façon automatique voire semi-automatique

térature (Del Mondo, 2011) de manière à manipuler des informations imparfaites. Ce modèle de graphe représente des relations de filiation entre observations géohistoriques, permettant ainsi d'en déduire les transformations existant entre elles. Nous proposons ensuite une approche semi-automatique permettant de peupler un tel graphe à partir des observations stockées dans la base de données spatiale et temporelle. Cette approche consiste à identifier les relations de filiation entre les observations géohistoriques issues d'un ensemble de sources géohistoriques. Dans ce but, nous proposons une approche en deux temps, constituée d'un algorithme inspiré des méthodes d'appariement de données géographiques et d'un algorithme d'identification de la nature des relations entre deux observations préalablement appariées. Afin de modéliser les imperfections des observations et des connaissances historiques, ces deux algorithmes s'appuient sur la théorie des fonctions de croyance (Shafer, 1976).

Contributions

Nous décrivons ici les contributions de cette thèse selon le domaine principal qu'elles concernent. Nous présentons tout d'abord les contributions en histoire, ou situées à l'intersection entre la géomatique et l'histoire. Nous détaillons ensuite les contributions concernant les données et modèles spatio-temporels, pour finir par les apports en géomatique.

Une lecture de l'évolution de l'administration urbaine de Paris au travers des grands plans de la ville

Une première contribution de cette thèse porte sur la mise en lumière, par les plans topographiques, de la mutation qui affecte l'administration de Paris après la Révolution française. La centralisation progressive des outils de gestion de la ville, qui commence dès la création de la préfecture de la Seine et se termine avec l'arrivée du baron Haussmann est un processus déjà connu des historiens. Les grands plans de Paris utilisés au sein de la préfecture sont des fenêtres sur l'intérieur de l'administration qui ont été peu exploitées. La réalisation de l'atlas de Jacoubet aux environs de 1830, que nous explorons en détail dans cette thèse, révèle en partie les dessous de l'évolution de l'administration qui mèneront aux grands travaux des années 1860.

Explorer et critiquer le contenu des sources cartographiques grâce au géoréférencement

L'existence de tracés géodésiques (méridienne, etc.) dans les plans topographiques de Paris offre la possibilité de géoréférencer ces plans sans utiliser de repères placés sur des objets géographiques inchangés comme cela est généralement fait. En utilisant de tels points sur deux thèmes cartographiques différents de l'atlas parisien réalisé par Edmé Verniquet à la fin du XVIII^e siècle, nous pouvons analyser les distorsions géométriques de chacun de ces thèmes par rapport au reste de la carte. Ces mesures permettent alors d'explorer certaines hypothèses sur la phase de levé topographique de l'atlas posées lors d'une analyse qualitative. Elles nous permettent en particulier de faire apparaître les différentes phases de levé topographique de l'atlas.

Plus généralement, nous montrons comment le géoréférencement peut être un outil d'analyse pertinent d'un plan et compléter une première étude historique.

Un modèle de graphe géohistorique adapté aux données imparfaites

Les observations géohistoriques sont imparfaites spatialement et temporellement : leur localisation spatiale est erronée en raison des erreurs planimétriques des sources, leur forme modifiée par les choix de représentation et leur localisation temporelle est incertaine. Peu de modèles spatio-temporels formels permettent de prendre en compte ces diverses imperfections, et nous n'en connaissons pas qui soient adaptés à la modélisation d'objets géographiques imparfaits en transformation dans le temps. L'extension proposée d'un modèle de graphe spatio-temporel de la littérature, proposé par Del Mondo (2011), offre la possibilité d'intégrer de telles informations dans un modèle formel pouvant être instancié dans une base de données relationnelle. Le graphe étendu que nous proposons est nommé *graphe géohistorique* afin de le distinguer plus aisément du modèle sur lequel il s'appuie.

Un processus d'intégration semi-automatique de données spatio-temporelles imparfaites

L'intégration de données dans un modèle spatio-temporel est rarement abordée, les données devant être préalablement traitées pour être représentables dans un modèle précis. L'approche d'appariement de données spatio-temporelles proposée dans cette thèse est un processus permettant d'intégrer des données géohistoriques dans un modèle de graphe spatio-temporel de façon semi-automatique. En outre, ce processus prend en compte certaines imperfections des données, inévitables lorsqu'il s'agit de données historiques.

Une modélisation de l'identité d'objets spatio-temporels dans le cas de connaissances incertaines

Les identités des objets spatio-temporels sont généralement considérées comme connues *a priori* au sein des modèles existants. Or, cela implique une connaissance fine des objets spatio-temporels que nous ne possédons pas dans notre cas, ces objets étant formés d'une chaîne d'observations géohistoriques. L'expression de l'identité sous la forme d'un ensemble d'attributs communs à ces observations et d'un ensemble de mesures de similarités entre attributs permet de traiter l'identité des objets au sein même de la base et non en amont. Cela permet d'alléger sensiblement la phase de collecte des données pour déplacer celle de détection des observations de même identité après le stockage des données dans la base. De plus, cela augmente la robustesse de la base de données à des données hétérogènes.

Une base de données spatio-temporelle extensible et adaptable à différents points de vue

L'approche adoptée pour le peuplement d'un graphe géohistorique à partir d'observations est divisée en deux phases, une première permettant de découvrir les observations dont les identités sont liées, et une seconde raffinant le résultat de la première pour identifier finement le type de relation selon que les observations décrivent la même entité ou des entités liées par un processus de transformation. La

première phase est coûteuse en temps de calcul mais nécessite une définition de l'identité des observations moins précise que la seconde phase. Cette dernière peut être réalisée à la volée pour un usage donné, selon le point de vue d'un expert historien. Le graphe finalement produit peut donc varier d'un expert à l'autre. Ainsi, le modèle proposé est adaptable à plusieurs usages sans qu'il soit nécessaire de modifier la base de données.

Une approche d'appariement de données géographiques générique et multi-source

L'approche de peuplement d'un graphe géohistorique est fondée sur un processus d'appariement de données géographiques multicritère. Ce processus permet un appariement entre N couches de données simultanément. L'approche, fondée sur la théorie des fonctions de croyances et étendant la proposition de (Olteanu, 2008), est entièrement générique et permet de modéliser des connaissances imparfaites. De plus, cette approche permet en une seule passe la détection de liens d'appariement de toutes cardinalités ($1 : 0$ à $N : M$). Il s'agit donc, au delà d'un processus de construction d'un graphe spatio-temporel, d'une méthode d'appariement de données géographiques étendant les propositions existant dans la littérature.

Perspectives

Dans ce travail, nous avons tenté de couvrir l'intégralité d'un processus d'intégration de données historiques permettant de former des bases de données spatio-temporelles utilisables pour l'étude historique. Cependant, de nombreux points n'ont été abordés que partiellement et ouvrent des perspectives de recherche variées. Les perspectives concernant le cas précis du processus d'intégration semi-automatique d'observations géohistoriques sont abordées au cours de la conclusion du chapitre 6 (cf. section 6.5).

Requêtes spatio-temporelles sur les graphes géohistoriques construits

Le travail que nous présentons s'arrête à la construction de graphes géohistoriques mais n'aborde pas l'analyse des transformations de l'espace que permettent ces graphes, ce qui constitue pourtant une des finalités d'un SIG spatio-temporel. Une première façon d'effectuer cette analyse consiste à donner la possibilité d'effectuer des requêtes spatio-temporelles sur la base de données produite. Le modèle de graphe de Del Mondo (2011) sur lequel s'appuie notre proposition est instanciable dans une base de données relationnelle et s'accompagne d'un ensemble de fonctionnalités permettant de parcourir ce graphes et d'effectuer des requêtes complexes. Bien que nous ayons proposé dans le chapitre 5 plusieurs prédicats permettant de détecter certains types de processus spatio-temporels propres aux données urbaines, nous n'apportons pas de réponse quant au stockage dans une base de données du modèle de graphe étendu proposé. Deux pistes en particulier nous semblent pertinentes pour permettre d'effectuer des requêtes sur le graphe produit. Tout d'abord, une possibilité est de transformer le graphe géohistorique vers un graphe spatio-temporel selon le modèle de (Del Mondo, 2011). En effet, le modèle de graphe géohistorique est principalement nécessaire dans une optique de construction, lorsque les identités

des entités géographiques sous-jacentes sont inconnues. Transformé, le graphe géo-historique peut être alors instancié dans une base de données relationnelle et devient compatible avec les outils d'analyse proposés par Del Mondo. Une thèse actuellement en cours au laboratoire COGIT (Costes *et al.*, 2015) propose une extension aux données imparfaites d'un modèle de graphe temporel agrégé (George et Shekhar, 2006). Ce graphe est une structure de données fusionnées permettant de représenter efficacement les transformations d'un réseau dans le temps. En outre, il est possible d'effectuer des requêtes sur ce graphe sous la forme d'expressions régulières afin de détecter des processus spatio-temporels. Ce type de graphe stockant des données fusionnées, il se place explicitement après une étape d'appariement des différents états des réseaux. Notre approche procédant à un tel appariement, une piste est donc d'intégrer nos résultats à ce graphe. Toutefois, cette intégration nécessite des opérations de fusion des observations géohistoriques.

Outils d'exploration visuelle de graphes spatio-temporels

Nous n'avons pas abordé les questions d'exploration visuelle des graphes spatio-temporels. Les observations géohistoriques sont des objets complexes dont plusieurs aspects sont imparfaitement connus (temporalité, localisation, etc.). Le graphe géo-historique est également une structure de données complexe, difficile à représenter visuellement. Les visualisations simples du chapitre 6 sont très peu intuitives et ne sont adaptées qu'à de petits graphes peu denses. Pourtant, un des intérêts des graphes construits est de permettre l'exploration des transformations de l'espace à partir d'une définition de l'identité des observations, mais également la découverte des cas de sauts temporels ou d'inversion dans l'ordre des sources géohistoriques. De plus, les graphes géohistoriques créés peuvent être de grande taille.

L'exploration visuelle de ces graphes est cruciale pour qu'ils puissent être utilisés par des historiens pour analyser les transformations de l'espace urbain. De tels outils d'exploration sont encore à définir. Les propositions existantes de visualisation de données spatio-temporelles (Kraak, 1999; Hägerstrand, 1970) sont en l'état peu adaptées à décrire les transformations d'objets géographiques autres que des *moving objects*. Précisément sur ce type de données, Stefani (2010) propose un outil d'exploration visuelle des transformations de parties de monuments archéologiques fondé sur une co-visualisation d'un graphe spatio-temporel et un modèle en trois dimensions. Ce type de modèle est toutefois adapté seulement dans le cas de graphes de faible densité et peu complexes. Plus globalement, l'exploration de données spatio-temporelles complexes est un problème encore ouvert qu'il serait intéressant d'aborder dans le cadre de la production de graphes géohistoriques. Sur ce thème, des travaux préliminaires réalisés au laboratoire COGIT lors d'un stage de master 2 ont permis d'explorer la cartographie de bâtis en transformation représentés par plusieurs états. Cette exploration a été réalisée à l'aide d'un modèle d'optimisation tentant de maximiser la lisibilité de la carte résultante à partir d'un jeu de règles de sémiologie graphique portant sur la couleur et la symbolisation du contour et du remplissage des bâtis. Bien que les pistes ouvertes par ce travail restent à approfondir, il a mis en évidence l'intérêt de la cartographie 2D pour l'exploration visuelle de données spatio-temporelles. Il a aussi pointé les limites de cette approche lorsque le nombre d'états successifs est important, ou que la configuration des transformations

est complexe.

Mise à jour

Un graphe géohistorique a vocation à évoluer sur plusieurs niveaux. Tout d’abord, il doit être capable d’intégrer de nouvelles observations et de nouvelles sources géohistoriques. On peut également se poser la question de l’ajout d’informations par des experts et de la modification d’un graphe existant. Il peut s’agir de supprimer d’éventuelles erreurs, d’éditer manuellement le graphe ou encore de renforcer la fiabilité de certaines relations. Le processus de construction d’un graphe géohistorique devient collaboratif, l’approche semi-automatique devenant alors soit un support pour l’expertise, soit un expert comme un autre. Ceci pose un certain nombre de questions ouvertes :

- comment intégrer et réconcilier des modifications concurrentes du graphe ? Quelles procédures de validation du graphes et de ses modifications adopter ?
- quel degré de fiabilité accorder aux différents utilisateurs d’un même graphe ? De la même façon, quelle place donner à la méthode de construction semi-automatique ?

De la même manière se pose la question de la mise à jour uniquement par la méthode semi-automatique. Par exemple, comment intégrer des modifications d’un graphe par une exécution de la méthode avec des paramètres différents de ceux ayant servi à sa création initiale ? Les modifications sont-elles même compatibles ou s’agit-il en réalité d’un nouveau graphe avec une vision différente de l’identité ?

Plus globalement, toutes ces questions dirigent la recherche vers la mise en place d’un modèle collaboratif de construction de graphes géohistoriques. La saisie et l’édition collaborative de bases de données historiques est actuellement au cœur du travail du groupe GeoHistoricalData⁴² rassemblant des chercheurs de différentes disciplines et établissements⁴³. Ce groupe de travail vise la vectorisation de cartes anciennes sur la France entière pour en étudier les dynamiques depuis les XVII-XVIII^e siècles. Actuellement, la majeure partie de la carte de Cassini a été vectorisée (cf. Perret *et al.* (2015)). Notre travail devrait être intégré au sein de ce groupe de travail afin de créer une base de données spatio-temporelles en graphe sur le territoire français. Les questions que nous posons ici seront alors nécessairement abordées.

Intégration de sources de données hétérogènes

Dans cette thèse, nous nous sommes limité au cas des plans topographiques de Paris à une époque où la précision géométrique est au cœur du travail de cartographe. De telles sources de données sont toutefois rares et ne constituent qu’une petite partie des sources géohistoriques utilisables pour analyser les transformations de l’espace ancien (voir à ce propos le chapitre 1 qui présente la diversité de sources utilisables). Dans un premier temps, d’autres types de sources cartographiques pourraient être considérés, notamment les plans plus anciens et moins exacts ou encore les plans en perspective. Cela soulève alors des questions de robustesse de notre approche. En effet, les données que nous utilisons sont de nature proche des données géographiques 2D actuelles et sont relativement homogènes. L’intégration de données

42. <https://www.geohistoricaldata.org/>

43. EHESS, IGN, CEA, IFSTTAR, BnF notamment.

d'autres natures - par exemple des bâtiments en perspective - nécessite d'étendre la proposition en ajoutant des étapes de prétraitement visant à harmoniser les données. Une approche plus souple serait d'intégrer au sein même du processus d'appariement spatio-temporel des méthodes d'harmonisation des données ou encore d'utiliser des mesures adaptées à des données de natures différentes.

Analyse quantitative des relations d'héritage entre plans topographiques de Paris

Dans le chapitre 3, nous montrons comment le géoréférencement peut être utilisé pour analyser en finesse le contenu d'un plan topographique. Dans le chapitre 2, nous posons un certain nombre d'hypothèses sur les relations d'héritage entre les différents plans de Paris, notamment entre les atlas de Jacoubet et de Verniquet. Cependant, nous ne parvenons pas à vérifier réellement ces hypothèses par manque de sources. Au delà de l'analyse de chaque plan indépendamment, les mesures de distorsions rendues possibles par le géoréférencement peuvent être utilisées pour comparer les plans entre eux. En particulier, la recherche de corrélations spatiales entre distorsions sur plusieurs plans peuvent donner des indices sur la réutilisation des relevés de triangulation.

Annexes

Annexe A

Carte des points de triangulation de l'atlas de Verniquet

Légende

- Points de station principaux
- Monuments ayant servi à construire les triangles
- Points de stations ajoutés ultérieurement
- Points de station secondaires rapportés dans [Pronteau,1986]
- Points de station secondaires supplémentaires

Annexe B

Atlas National de la Ville de Paris par Edmé
Verniquet, 1783-1791 -Géoréférencement

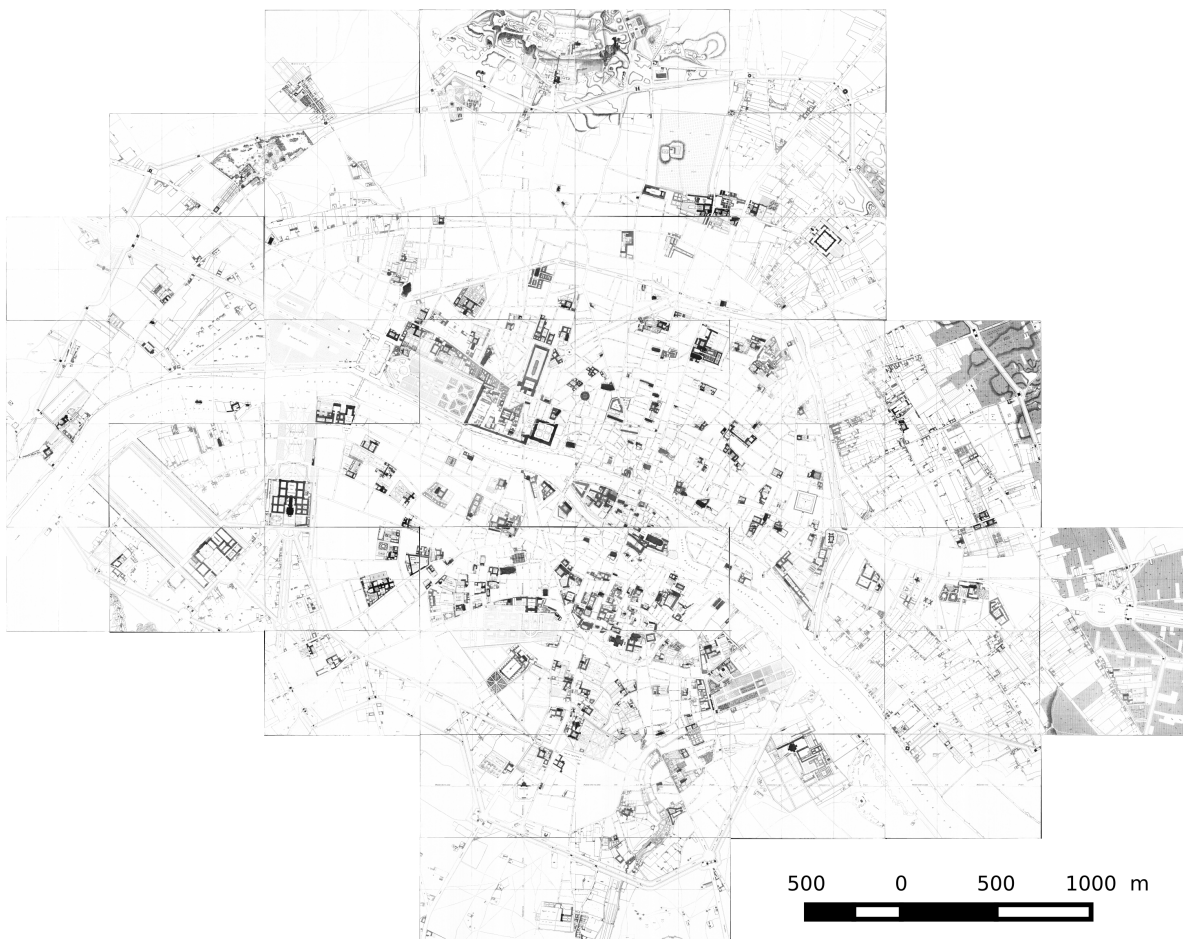


FIGURE B.1 – *Atlas National de la Ville de Paris* par Edmé Verniquet, fac-similé en 72 planches de Marcel Legoux daté de 1983. Assemblage après géoréférencement.

Annexe C

La topographie de Paris, 1803-1808 par Nicolas
Maire - Géoréférencement



FIGURE C.1 – *La topographie de Paris ou plan détaillé de la ville de Paris et de ses faubourgs* par Nicolas Maire, édition de 1808 en 20 feuilles, partie de l'atlas du même nom. Assemblage après géoréférencement.

Annexe D

Atlas Général de la Ville et des Faubourgs de Paris,
1825-1836 par Théodore Simon Jacoubet
-Géoréférencement

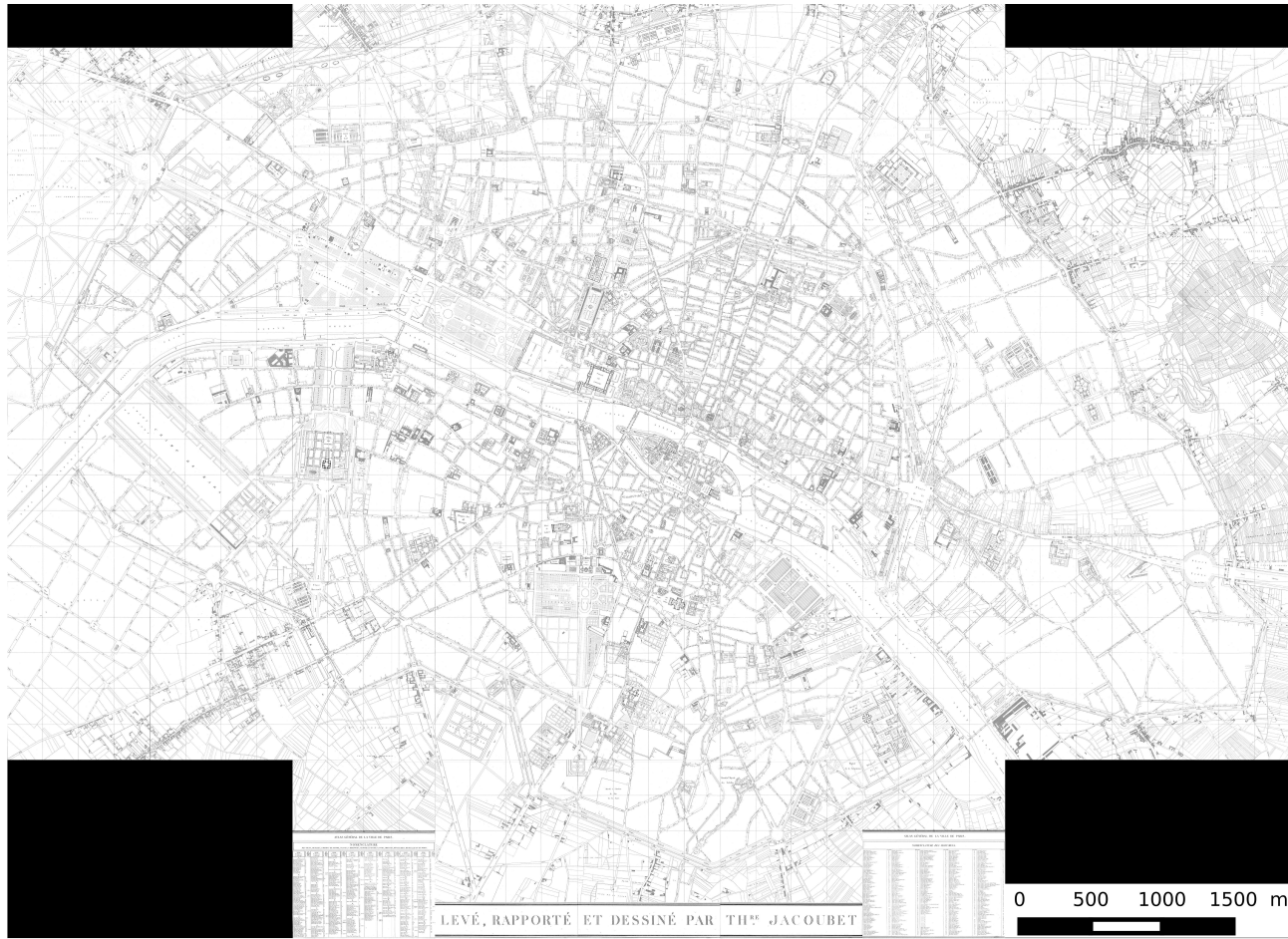


FIGURE D.1 – *Atlas Général de la Ville et des Faubourgs de Paris* par Théodore Jacobet, édition de 1836 en 54 planches. Les planches 1 et 2 (titre), 8 et 9 (plan d'assemblage), 46-47 et 53-54 (nomenclatures et statistiques), non cartographiques, ne sont pas affichées. Assemblage après géoréférencement.

Annexe E

Atlas municipal des vingt arrondissements de la ville de Paris, 1888 -Géoréférencement



FIGURE E.1 – Atlas de Paris en 16 feuilles, édition de 1888, réalisé par les géomètres du Service du Plan de Paris, assemblé après géoréférencement des feuilles.

Annexe F

Agrandissements des réseaux de rue vectorisés



FIGURE F.1 – Réseau viaire vectorisé de l’atlas de Verniquet superposé au réseau de Paris actuel issu de la BDTPO.



FIGURE F.2 – Réseau viaire vectorisé de l’atlas de Vasserot superposé au réseau de Paris actuel issu de la BDTPO.



FIGURE F.3 – Réseau viaire vectorisé de l’atlas de Jacoubet superposé au réseau de Paris actuel issu de la BDTPOPO.



FIGURE F.4 – Réseau viaire vectorisé de l’atlas municipal superposé au réseau de Paris actuel issu de la BDTPOPO.

Annexe G

Indices temporels des planches de l'Atlas de Jacoubet

I	3	◀ 1828 : Construction de l'église Sainte-Marie des Batignolles ◀ 1830 : Création de la commune Batignolles-Monceau.
	10	▶ 1825 : Projet de lotissement de la plaine de Passy
	11	▶ 1825 : Percement des rues de Byron et de Châteaubriand ▶ 4 novembre 1829 : Ordonnance de percement de la rue Fortin. ◀ 18 mars 1836 : Ordonnance de percement des rues de Newton et des Bassins.
	12	▶ 2 février 1826 : Ordonnance de percement des rues du quartier Europe ◀ 27 septembre 1836 : Ordonnance d'alignement de la rue Duras, Petite rue Verte et Roquepine.
II	4	▶ 2 février 1826 : Ordonnance d'ouverture des rues du Quartier Europe et de la rue Fontaine. ◀ 1836 : Lotissement de la rue Fontaine
	5	▶ 27 septembre 1826 : Ordonnance d'ouverture de la rue Frochot ▶ 2 février et 16 août 1825 : Ordonnance d'ouverture et de nommage de la rue du Delta ◀ 23 août 1833 : ordonnance d'alignement de la ruelle Beauregard ou rue Beauregard des Martyrs.
	6	▶ 27 septembre 1826 : Ordonnance d'ouverture de la rue Frochot ▶ 21 janvier 1827 : Ordonnance d'ouverture de la rue d'Hauteville entre la rue du Chevet-de-l'Église et le boulevard extérieur. ▶ 29 novembre 1826 : Ordonnance d'élargissement de la rue du Château Landon à 15 mètres.
	13	▶ 27 septembre 1826 : Ordonnance d'ouverture de la rue Frochot ▶ 21 janvier 1827 : Ordonnance d'ouverture de la rue d'Hauteville entre la rue du Chevet-de-l'Église et le boulevard extérieur. ◀ 1831 : Ouverture sans autorisation de la rue de Milan. ◀ 1829-1830 : Construction de la cité d'Antin
		▶ 27 septembre 1826 : Ordonnance d'ouverture de la rue Frochot ▶ 7 octobre 1830 : Ordonnance d'ouverture de la rue Navarin.

		<ul style="list-style-type: none"> ◀ 1831 : Ouverture sans autorisation de la rue de Milan. ▶ 1830 : Percement de la rue Vivienne suivant l'ordonnance du 17 janvier 1830. ◀ 1832 : Ouverture sans autorisation du passage Laferrière.
	15	<ul style="list-style-type: none"> ▶ 16 décembre 1829 : Ordonnance autorisant la transformation du passage du Grand Saint-Michel en rue. ▶ 1827 : Ouverture des rues environnant l'église Saint-Vincent-de-Paul. ◀ 1836 : Fin de l'élargissement de l'impasse des Filles-Dieu à 8 mètres. ◀ 1832 : Suppression de l'impasse des Babillards.
III	7	▶ 1830 : La rue Lafayette est nommée comme telle suite à la révolution de 1830. De même pour les quai Jemmapes et de Valmy.
	16	<ul style="list-style-type: none"> ▶ 1 mars 1826 : Ordonnance autorisant le percement de la rue du Canal Saint Martin. ▶ 6 juin 1825 : ordonnance d'ouverture des rues Claude Villefaux/Vellefosse/Villefosse et Châstillon. ▶ 6 décembre 1827 : ordonnance d'élargissement de l'impasse Saint Louis (rue Alibert). ▶ 1830 : Rue du corbeau nommée officiellement. ◀ 1836 : Fin des travaux de l'avenue Richerand.
	17	
	18	
IV	19	▶ 6 avril 1832 : Ordonnance d'élargissement de la rue du chemin de Versailles à 13 mètres.
	20	<ul style="list-style-type: none"> ▶ 1832 : Ouverture de la rue Malar de la rue de la Triperie au quai d'Orsay. ▶ 1833 : Ouverture de la rue de la Vierge entre la rue de l'Université et le quai d'Orsay.
	21	<ul style="list-style-type: none"> ▶ 1829 : Fin de la construction du pont des Invalides. ▶ 1828 : Percement de la place de Bellechasse. ▶ 1830 : Le pont de Louis XVI est renommé pont de la Concorde, de même pour la place homonyme ◀ 25 décembre 1836 : Installation de l'obélisque de la place de la Concorde. ◀ 7 mars 1827 : ordonnance d'élargissement des rues de Bourgogne, de Coutry, Saint-Dominique .
	28	<ul style="list-style-type: none"> ▶ 1826 : La gare fluviale de Grenelle est terminée. ◀ 22 octobre 1830 : La commune de Grenelle est fondée et reçoit son nom officiel.
	29	▶ 1825 : Projet de lotissement de Grenelle.
	30	<ul style="list-style-type: none"> ▶ 1828 : La place Bellechasse et les rues Casimir Perier, de Champigny et Martignac sont ouvertes ◀ 18 mai 1838 : Ordonnance de percement de la rue Barbet de Jouy.

V	22	<ul style="list-style-type: none"> ▶ 1832 : La rue Louis-Philippe 1e est renommée rue d'Alger après son percement en 1830. ▶ 1828 : Percement de la place de Bellechasse. ◀ 1834 : Fin de la construction du pont des Saint-Pères.
	23	<ul style="list-style-type: none"> ▶ 2 juillet 1833 : Percement et nommage de la rue de la Bourse. ◀ 6 mai 1836 : Ordonnance d'élargissement de la rue du Renard. ◀ 20 mai 1836 : Ordonnance de percement de la rue Sainte-Opportune.
	24	<ul style="list-style-type: none"> ▶ 6 mars 1828 : Ordonnance d'élargissement à 13 mètres de la rue Neuve Saint-Nicolas. ▶ 27 mai 1829 : Ordonnance de percement de la rue Neuve Bourg-l'Abbé. ◀ 14 janvier 1829 : Ordonnance d'élargissement à 10 mètres des rues Transnonain et des Vertus. ◀ 16 mai 1833 : Ordonnance d'élargissement à 12 mètres des rue Aumaire, des Gravilliers, Jean-Robert, Pastourelle, Michel le Comte , des Vieilles Andriettes, du Maure, etc. ◀ 21 juin 1826 : Ordonnance d'alignement des rues Cléry, Saint-Apolline, Neuve Saint-Denis, Beauregard et Tracy
	31	<ul style="list-style-type: none"> ▶ 1834 : Le pont des Saint-Pères est terminé ▶ 1831 : l'hôtel Brochet de Saint Prest devient la direction des Mines et des Ponts et Chaussées.
	32	<ul style="list-style-type: none"> ▶ 1831 : La rue de l'évêché de la Cité est supprimée. ▶ 1835 : La rue Racine est en cours de percement. ◀ 1837 : Ordonnances de percement de la rue d'Arcole et de Constantine.
	33	<ul style="list-style-type: none"> ▶ 20 juillet 1830 : Ordonnance d'élargissement de la rue du Lion. ▶ 1834 : Le pont Louis-Philippe est terminé. ▶ 18 août 1833 : ordonnance de percement de la rue du Pont Louis-Philippe. ◀ 1836 : Percement de la rue Lobau.
VI	25	<ul style="list-style-type: none"> ▶ 9 septembre 1829 : Construction du Marché de Popincourt. ◀ 31 mars 1835 : Ordonnance d'alignement de la rue de Normandie. ◀ 1829 : Ouverture du passage Crussol.
	26	<ul style="list-style-type: none"> ▶ 1826 : Projet de construction de la prison de la Grande Roquette ◀ 1830 : Inauguration de la prison de la Grande Roquette.
	27	
	34	<ul style="list-style-type: none"> ▶ 8 juin 1834 : Ordonnance d'alignement des rues des Douze-Portes, Neuve Saint-Pierre et Saint-Gilles ▶ 15 juillet 1829 : Ordonnance de percement de la rue Neuve Lappe. ▶ 25 février 1829 : Ordonnance de percement de la rue de l'Orme entre la Cour du Salpêtre et la rue Saint-Antoine. ▶ 27 mai 1827 : Ordonnance de percement de la partie sud de la rue de la Planchette. ◀ 6 mai 1827 : Ordonnance d'alignement des rues de Charonne et de la Roquette.

VII		<ul style="list-style-type: none"> ◀ 27 janvier 1831 : La rue de Lappe est renommée rue Louis-Philippe. ◀ 1 mai 1828 : Ordonnance d'alignement des rues de Charenton, Terres-Fortes et Moreau.
	35	<ul style="list-style-type: none"> ▶ 1830 : Début du projet de la prison de la Petite Roquette. ◀ 6 mai 1827 : Ordonnance d'élargissement de la rue de la Muette à 13 mètres et de la rue des Murs de la Roquette.
	36	◀ 6 mai 1827 : Ordonnance d'élargissement de la rue de Charonne et de l'avenue des Ormeaux.
	37	▶ 1829 : Construction de l'église de Grenelle.
	38	▶ 1825 : Projet de lotissement de Grenelle.
	39	◀ 18 mai 1838 : Ordonnance de percement du boulevard de Vaugirard entre la rue de Sèvre et la rue des Fourneaux.
	48	◀ 1838 : Ouverture de la rue Mouton-Duvernet.
VIII	40	<ul style="list-style-type: none"> ▶ 21 septembre 1827 : Ordonnance d'élargissement à 12 mètres de la rue Campagne-Première. ◀ 1835 : Élargissement de la rue Campagne-Première suivant l'ordonnance du 21 septembre 1827.
	41	<ul style="list-style-type: none"> ▶ 11 décembre 1830 : Loi rendant son nom et fonction au panthéon. ▶ 9 août 1828 : Ordonnance de percement de la rue Soufflot de la rue Saint-Jacques au palais du Luxembourg. ▶ 6 mai 1827 : Ordonnance de percement de la rue Pascal.
	42	<ul style="list-style-type: none"> ▶ 1832 : Construction de l'amphithéâtre d'anatomie de Clamart. ◀ 1838 : Construction de la place Saint-Victor.
	49	◀ 1838 : Ouverture de la rue Mouton-Duvernet.
	50	<ul style="list-style-type: none"> ▶ 1832 : Inauguration du dépôt de mendicité de Lourcine. ◀ 1836 : Inauguration de l'hôpital de Lourcine.
	51	▶ 2 décembre 1829 : Ordonnance d'élargissement à 23 mètres de la rue du Petit Gentilly.
IX	43	<ul style="list-style-type: none"> ▶ 1 juin 1828 : Ordonnance d'alignement des rues des Charbonniers, de Bercy et Lacuée. ▶ 9 août 1832 : Construction du pont de la Gare (de Bercy). ▶ 1830 : Le pont du Roi reprend le nom de pont d'Austerlitz. ◀ 1839 : Construction de la première gare ferroviaire d'Austerlitz.
	44	▶ 1830 : Élargissement de la rue des Quatre-Chemins.
	45	▶ 1830 : Élargissement de la rue des Quatre-Chemins.
	52	

Annexe H

Évaluation des méthodes de rangement de quantités floues par rapport à un ordre de référence

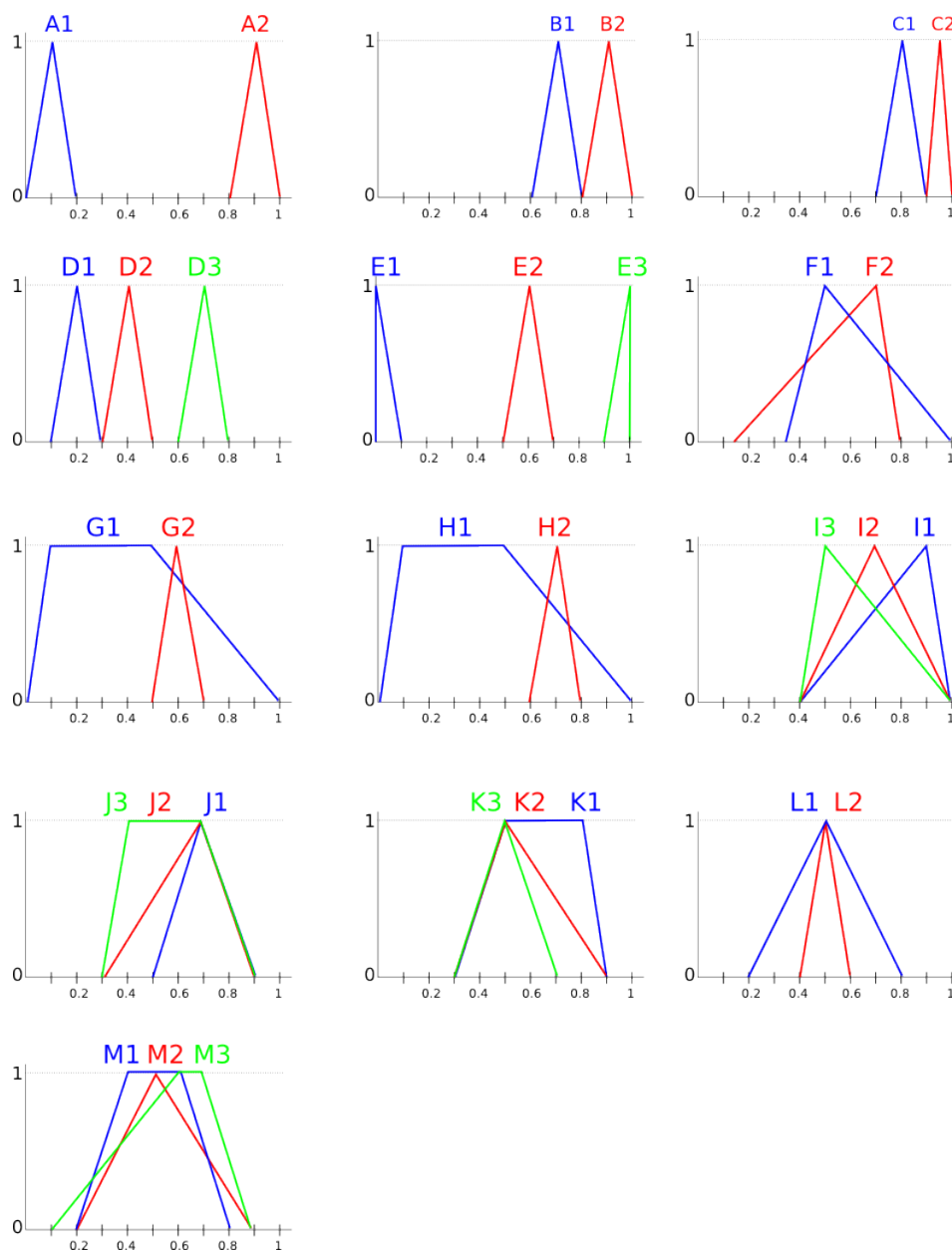


FIGURE H.1 – Configurations de quantités floues évaluées. Exemples d'après (Bortolan et Degani, 1985) et (Chen et Lu, 2001b).

TABLE H.1 – Évaluation des méthodes de tri d’ensembles flous sur les exemples en figure H.1. Les lignes en jaune aboutissent au même rangement que celui de référence quelque soit la situation. Les cases en gris identifient le premier cas d’échec du tri par rapport à l’ordre de référence. Partie 1.

	A	B	C	D	E	F	G
Référence	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ ~ ■	■ λ ■
Yager (1981)F1	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
F2	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
F3	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ ~ ■	■ λ ■
Chang (1981)	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
Adamo (1980)0.9M	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
0.5	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
Baas et Kwakernaak (1977)	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
Baldwin et Guild (1979b)l.p	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
r	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
r.a	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
Kerre (1982)	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
Jain (1977)K=1	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
K=2	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
K=0.5	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
Dubois et Prade (1983)PD	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
PSD	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
ND	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
NSD	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
Kim et Park (1990)K=1	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
K=0.5	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
K=0	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
Fortemps et Roubens (1996)	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ ~ ■	■ λ ■
Liou et Wang (1992) $\gamma = 1$	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
$\gamma = 0.5$	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ ~ ■	■ λ ■
$\gamma = 0$	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
Abbasbandy et Hajjari (2009)	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
Asady (2011)	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ ~ ■	■ λ ■
Choobineh et Li (1993)	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ ~ ■	■ λ ■
Abbasbandy <i>et al.</i> (2013)p=1	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
p=2	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
p=3	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
p=4	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■
p=5	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■	■ λ ■

TABLE H.3 – Évaluation des méthodes de tri d'ensembles flous sur les exemples en figure H.1. Les lignes en jaune aboutissent au même rangement que celui de référence quelque soit la situation. Les cases en gris identifient le premier cas d'échec du tri par rapport à l'ordre de référence. Partie 2.

	H	I	J	K	L	M
Référence	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
Yager (1981)F1	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
F2	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
F3	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
Chang (1981)	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
Adamo (1980) 0.9M	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
0.5	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
Baas et Kwakernaak (1977)	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
Baldwin et Guild (1979b)l.p	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
r	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
r.a	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
Kerre (1982)	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
Jain (1977)K=1	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
K=2	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
K=0.5	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
Dubois et Prade (1983)PD	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
PSD	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
ND	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
NSD	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
Kim et Park (1990)K=1	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
K=0.5	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
K=0	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
Fortemps et Roubens (1996)	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
Liou et Wang (1992) $\gamma = 1$	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
$\gamma = 0.5$	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
$\gamma = 0$	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
Abbasbandy et Hajjari (2009)	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
Asady (2011)	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
Choobineh et Li (1993)	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
Abbasbandy <i>et al.</i> (2013)p=1	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
p=2	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
p=3	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
p=4	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■
p=5	■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■ λ ■	■ λ ■	■ λ ■ λ ■

Annexe I

Évaluation des méthodes mesurant de degré d'antécédence de deux temporalités floues

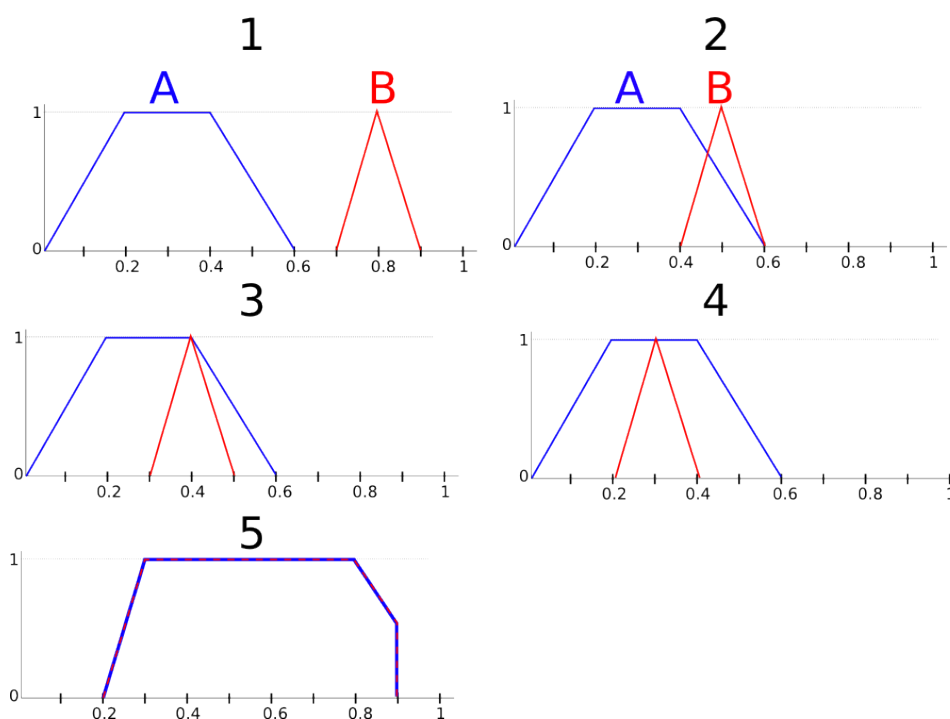


FIGURE I.1 – Configurations de temporalités floues sur lesquelles sont évaluées les méthodes calculant un degré d'antécédence.

TABLE I.1 – Degré d'antécédence des temporalité floues représentées dans la figure I.1 avec différentes méthodes fondées sur la dominance entre sous-ensembles flous

	1		2		3		4		5	
	ant(A,B)	ant(B,A)
(De Runz, 2008)	1	0	1	0	0.83	0.16	0.5	0.5	1	1
(Kołodziejczyk, 1986) K1	1	0	0.86	0.14	0.7	0.3	0.5	0.5	0.5	0.5
. K2	1	0	0.86	0.14	0.7	0.3	0.5	0.5	0.5	0.5
(Chen et Lu, 2001b) $\beta=1$	-0.35	0.35	-0.05	0.05	0.05	-0.05	0.15	-0.15	0	0
. $\beta=0.5$	-0.5	0.5	-0.2	0.2	-0.01	0.01	0	0	0	0
. $\beta=0$	-0.65	0.65	-0.35	0.35	-0.3	0.3	-0.15	0.15	0	0
1- (Dubois et Prade, 1983) PD	1	0	0.35	0	0	0	0	0	0	0
1- PSD	1	0	0.7	0.35	0.35	0.7	0	1	0.55	0.55
1- ND	1	0	1	0	1	0	1	0	0.5	0.5
1- NSD	1	0	1	0.65	1	1	1	1	1	1

Bibliographie

- ABADIE, N. (2012). *Formalisation, acquisition et mise en œuvre de connaissances pour l'intégration virtuelle de bases de données géographiques : Les spécifications au cœur du processus d'intégration*. Thèse de doctorat, Université Paris-Est. 241
- ABBASBANDY, S. et ASADY, B. (2006). Ranking of fuzzy numbers by sign distance. *Information Sciences*, 176(16):2405–2416. 202
- ABBASBANDY, S. et HAJJARI, T. (2009). A new approach for ranking of trapezoidal fuzzy numbers. *Computers & Mathematics with Applications*, 57(3):413–419. 202, 343, 344
- ABBASBANDY, S., NURAEI, R. et GHANBARI, M. (2013). Revision of sign distance method for ranking of fuzzy numbers. *Iranian Journal of Fuzzy Systems*, 10(4): 101–117. 202, 343, 344
- ADAMO, J. (1980). Fuzzy decision trees. *Fuzzy sets and systems*, 4(3):207–219. 201, 203, 343, 344
- ALIZAMIR, S., PARDALOS, P. M. et REBENNACK, S. (2008). Improving the neighborhood selection strategy in simulated annealing using the optimal stopping problem. In TAN, C. M., éditeur : *Simulated Annealing*, chapitre 18, pages 363–379. INTECH Open Access Publisher. 267
- ALLEN, J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843. 199
- ALPHAND, A. et FAUVE, M. (1888). Atlas municipal des vingt arrondissements de la ville de Paris. Dressé sous la direction de M. Alphand inspecteur général des ponts et chaussées, par les soins de M.L Fauve, géomètre en chef, avec le concours des géomètres du plan de Paris. BHVP, FOL AT 2. 78
- ALT, H. et GODAU, M. (1995). Computing the fréchet distance between two polygonal curves. *International Journal of Computational Geometry & Applications*, 5(01n02):75–91. 289
- APPRIOU, A. (1991). Probabilités et incertitude en fusion de données multi-senseurs. *Tiré à part- Office national d'études et de recherches aérospatiales*. 258, 259, 289
- ARCHIN, BERTHE, HACQ et PEYTIEUX (1825). Plan de Paris en 1825 avec le tracé des anciennes enceintes de cette ville. Dressé par Achin d'après le Plan de Verniquet, augmenté de tous les changemens survenus jusqu'à la présente époque. Gravé par Berthe. Écrit par Hacq. BNF, GE D-14458. 89

-
- ARNAUD, J.-L. (2008). *Analyse spatiale, cartographie et histoire urbaine*. Parenthèses, Marseille. 18, 19, 20, 22, 28, 30, 35, 59
- ASADY, B. (2011). Revision of distance minimization method for ranking of fuzzy numbers. *Applied Mathematical Modelling*, 35(3):1306–1313. 202, 203, 206, 343, 344
- ATKINSON, P., FOODY, G., DARBY, S. et WU, F. (2004). *GeoDynamics*. Taylor & Francis. 47
- AYERS, E. L. (2003). An overview : The differences slavery made : A close analysis of two american communities. *The American Historical Review*, 108(5):1299–1307. 51
- BAAS, S. M. et KWAKERNAAK, H. (1977). Rating and ranking of multiple-aspect alternatives using fuzzy sets. *Automatica*, 13(1):47–58. 201, 343, 344
- BAILLY, A. (1977). *La perception de l'espace urbain : les concepts, les méthodes d'études, leur utilisation dans la recherche urbanistique*. Centre de recherche d'Urbanisme. 32
- BAILLY, A. (1985). Distances et espaces, vingt ans de géographie des représentations. *L'Espace Géographique*, 3:197–205. 32, 47
- BAIOCCHI, V. et LELO, K. (2005). Georeferencing the historical maps of rome between the seventeenth and eighteenth centuries. *In CIPA 2005 XX International Symposium*. 123
- BALDWIN, J. et GUILD, N. (1979a). Comments on the fuzzy max operator of Dubois and Prade. *International Journal of Systems Science*, 10(9):1063–1064. 201
- BALDWIN, J. et GUILD, N. (1979b). Comparison of fuzzy sets on the same decision space. *Fuzzy sets and Systems*, 2(3):213–231. 201, 343, 344
- BALLETTI, C. (2000). Analytical and quantitative methods for the analysis of the geometrical content of historical cartography. *International Archives of Photogrammetry and Remote Sensing*, 33(B5/1 ; PART 5):30–37. 65, 113
- BALLETTI, C. (2006). Georeference in the analysis of the geometric content of early maps. *e-Perimetron*, 1(1):32–42. 120, 156
- BARBOT, M. (2013). La résidence comme appartenance. les catégories juridiques de l'inclusion sociale dans les villes italiennes sous l'ancien régime. *Histoire urbaine*, 36:29–48. 17
- BARTHELEMY, M., BORDIN, P., BERESTYCKI, H. et GRIBAUDI, M. (2013). Self-organization versus top-down planning in the evolution of a city. *Scientific Reports*, 3. 20, 172
- BATAILLON, M. (1950). La méditerranée et le monde méditerranéen à l'époque de philippe II. *Revue économique*, 1(2):232–241. 15

- BEERI, C., KANZA, Y., SAFRA, E. et SAGIV, Y. (2004). Object fusion in geographic information systems. *In Proceedings of the Thirtieth international conference on Very large data bases-Volume 30*, pages 816–827. VLDB Endowment. 252, 253
- BEN-AMEUR, W. (2004). Computing the initial temperature of simulated annealing. *Computational Optimization and Applications*, 29(3):369–385. 276, 292
- BENAVIDES, J. et KOSTER, E. (2006). Identifying surviving landmarks on historical maps. *e-Perimtron*, 1(3):194–208. 133
- BERMAN, M. L. (2009). Modeling and visualizing historical gis data. *In Proceedings of the Spatio-Temporal Workshop, Harvard University, Cambridge, Massachusetts*. 40
- BERMAN, M. L. (2013). Origins and development of the china historical gis. *Pelagios 3 Gazetteer Technical Workshop, Institute for the Study of the Ancient World*. 54, 370
- BERRY, B. J. L. (1964). Approaches to regional analysis : a synthesis. *Annals of the Association of American Geographers*, 54(1):2–11. 26
- BERTHAUT, H. M. A. (1898). *La carte de France, 1750-1898 : étude historique*, volume 1. Imprimerie du Service géographique. 134
- BERTY, A. et LENOIR, R. (188.). Plan archéologique de Paris : feuille XIV (université). Dressé par Albert Lenoir et Adolphe Berty. 182
- BODENHAMER, D. J., CORRIGAN, J. et HARRIS, T. (2010). *The Spatial Humanities : GIS and the Future of Humanities Scholarship*. Spatial humanities. Indiana University Press. 24
- BOL, P. (2007). The china historical geographic information system (CHGIS) : choices faced, lessons learned. *In Conference on Historical Maps and GIS, Nagoya University*. 53, 54, 58, 370
- BONNARDOT, A. (1851). *Études archéologiques sur les anciens plans de Paris des XVI^e, XVII^e et XVIII^e siècles*. Librairie Deflorenne. 69
- BORTOLAN, G. et DEGANI, R. (1985). A review of some methods for ranking fuzzy subsets. *Fuzzy sets and Systems*, 15(1):1–19. 201, 202, 203, 342, 375
- BOSSÉNO, C.-M. (1990). La guerre des estampes. circulation des images et des thèmes iconographiques dans l'Italie des années 1789-1799. *Mélanges de l'Ecole française de Rome. Italie et Méditerranée*, 102(2):367–400. 31
- BOUCHON-MEUNIER, B. (1995). *La Logique floue et ses applications*. Vie artificielle. Addison-Wesley France. 45, 46, 48, 49, 50, 190
- BOUDON, F. et BLÉCON, J. (1977). *Système de l'architecture urbaine : le quartier des Halles à Paris*. Atlas. Éditions du Centre national de la recherche scientifique. 19, 41

- BOURILLON, F. (2000). *Les Parisiens et la modernisation de la ville au XIXe siècle. Évaluer, transformer et construire la ville*. Thèse de doctorat, Université Paris 12-Val-de-Marne. Mémoire pour le diplôme d'Habilitation à diriger des recherches. 65
- BOUTIER, J. (2002). Mesures et triangulation de l'espace urbain. Le lever des plans de Paris à l'époque moderne (xvi^e-xviii^e siècles). *Le Monde des cartes*, 1(171):6–18. 119
- BOUTIER, J. (2007). *Les plans de Paris des origines, 1493, à la fin du XVIIIe siècle : étude, carto-bibliographie et catalogue collectif*. Bibliothèque Nationale de France-BNF. 69
- BOUTOURA, C. et LIVIERATOS, E. (2006). Some fundamentals for the study of the geometry of early maps by comparative methods. *e-Perimetron*, 1(1):60–70. 127, 130, 131, 151
- BOVE, B. (2012). Crise locale, crises nationales. rythmes et limites de la crise de la fin du Moyen Âge à Paris au miroir des prix fonciers. *Histoire Urbaine*, 33(1). 20
- BRAUDEL, F. (1951). La géographie face aux sciences humaines. *Annales. Économies, Sociétés, Civilisations*, 6(4):485–492. 15, 22
- BRETEZ, L. (1739). Plan de Paris dit "de Turgot". Réalisé de 1734 à 1739 sur une commande de Michel-Etienne Turgot, prévôt des marchands. 29, 142
- BROVELLI, M. A. et MINGHINI, M. (2012). Georeferencing old maps : a polynomial-based approach for como historical cadastres. *e-Perimetron*, 17(3). 137
- BROWN, L. G. (1992). A survey of image registration techniques. *ACM computing surveys (CSUR)*, 24(4):325–376. 122
- BUCHER, B. (2005). Décrire la généalogie de données géographiques. In *INFORSID*, pages 303–318. 180
- BURRUS, N. et LESAGE, D. (2007). Theory of evidence. Rapport technique, LRDE, EPITA. 282
- CASSELLE, P., éditeur (2000). *La Commission des embellissements de Paris ? Rapport à l'Empereur Napoléon III*. Numéro 23 de Cahier de la Rotonde. Paris : Rotonde de la Villette. 78, 105
- CASSINI DE THURY, C.-F. (1744a). Carte qui comprend tous (sic) les lieux de la France qui ont été (sic) déterminés par les opérations géométriques / par M. Cassini de Thury. 119
- CASSINI DE THURY, C.-F. (1744b). Nouvelle carte qui comprend les principaux triangles qui servent de fondement à la description géométrique de la France, levée par ordre du Roy par mess. Maraldi et Cassini de Thury. 119
- CHANG, W. (1981). Ranking of fuzzy utilities with triangular membership functions. In *Proc. Int. Conf. on Policy Anal. and Inf. Systems*, volume 263, page 272. 201, 202, 343, 344

- CHEN, L.-H. et LU, H.-W. (2001a). An approximate approach for ranking fuzzy numbers based on left and right dominance. *Computers & Mathematics with Applications*, 41(12):1589–1602. 201
- CHEN, L.-H. et LU, H.-W. (2001b). An approximate approach for ranking fuzzy numbers based on left and right dominance. *Computers & Mathematics with Applications*, 41(12):1589–1602. 202, 203, 205, 206, 342, 346, 375
- CHEN, S.-H. (1985). Ranking fuzzy numbers with maximizing set and minimizing set. *Fuzzy sets and Systems*, 17(2):113–129. 201, 203
- CHENG, C.-H. (1998). A new approach for ranking fuzzy numbers by distance method. *Fuzzy sets and systems*, 95(3):307–317. 202
- CHEVALIER, L. (1952). La formation de la population parisienne au XIX^e siècle. *Annales, Économies, Sociétés, Civilisations*, 7(1):93–95. 65
- CHOOBINEH, F. et LI, H. (1993). An index for ordering fuzzy numbers. *Fuzzy sets and Systems*, 54(3):287–294. 201, 203, 343, 344
- CHRISMAN, N. R. (1999). What does 'gis' mean? *Transactions in GIS*, 3(2):175–186. 39
- CHRYSSA PAPAOKOSTA, C. I. et VASSILAKI, D. (2009). The use of linear features as ground control information for the georeferencing of old aerial photos. *Journal of Historical Geography*, 30:1–21. 123
- CHU, T.-C. et TSAO, C.-T. (2002). Ranking fuzzy numbers with an area between the centroid point and original point. *Computers & Mathematics with Applications*, 43(1):111–117. 202
- CLARAMUNT, C. et THERIAULT, M. (1996). Toward semantics for modelling spatio-temporal processes within gis. *Advances in GIS Research I*, pages 27–43. 230
- CLARAMUNT, C. et THÉRIAULT, M. (1995). Managing time in GIS : an event-oriented approach. *In Recent Advances in Temporal Databases*, pages 23–42. Springer. 220
- CONSORTIUM, O. G. (2003a). OGC Reference Model. Rapport technique, Open Geospatial Consortium. 184
- CONSORTIUM, O. G. (2003b). Topic 20 : Observations and Measurements. Rapport technique, Open Geospatial Consortium. 177
- COSGROVE, D. E. (1984). *Social formation and symbolic landscape*. University of Wisconsin Press. 14
- COSTES, B. (2012). Appariement cassini - rge (géométries ponctuelles). Rapport technique, GeoPeuple. 43

- COSTES, B., PERRET, J., BUCHER, B. et GRIBAUDI, M. (2015). An aggregated graph to qualify historical spatial networks using temporal patterns detection. *In 18th AGILE International Conference on Geographic Information Science*. 249, 314
- COURVILLE, S. (1995). *Introduction à la géographie historique*. Géographie historique. Presses de l'Université Laval. 15
- COYECQUE, E. (1909). *Les plans cadastraux de la ville de Paris aux archives Nationales...* Daupeley-Gouverneur. 73
- DE CHABROL DE VOLVIC, G. J. G. (1823). *Recherches statistiques sur la ville de Paris et le département de la Seine : recueil de tableaux*. Numéro vol. 2 de Recherches statistiques sur la ville de Paris et le département de la Seine : recueil de tableaux. Imprimerie royale. 82, 83
- DE LANGE, N., PANAYOTOV, A. et REESE, G. (2013). Mapping the jewish communities of the byzantine empire. <http://www.byzantinejewry.net/>. Accessed : 2013-11-25. 52, 53, 57, 370
- DE L'ISLE, G. (1716). Le plan de Paris, ses Faubourgs et ses environs. USUEL-A-166. 29
- DE ROCHEFORT, J. (1672-1674). Le nouveau plan de Paris. Fac-similé du premier plan de Jouvin de Rochefort, réduction de 1/4. 37
- DE RUNZ, C. (2008). *Imperfection, temps et espace : modélisation, analyse et visualisation dans un SIG archéologique*. Thèse de doctorat, Université de Reims Champagne-Ardenne, France. 31, 45, 46, 48, 49, 50, 200, 202, 206, 249, 346, 370
- DE RUNZ, C., DESJARDIN, E., PIANTONI, F. et HERBIN, M. (2010). Anteriority index for managing fuzzy dates in archaeological gis. *Soft Computing*, 14(4):339–344. 203
- DE VASSALIEU, B. (1609). Portrait de la ville, cité et université de Paris, avec les faubours d'icelle. Faict par Vassalieu, Nicolay, topographe et ingénieur ordinaire de l'artillerie de France . BHVP, USUEL-A-103 b. 37
- DEL MONDO, G. (2011). *Un modèle de graphe spatio-temporel pour représenter l'évolution d'entités géographiques*. Thèse de doctorat, Université de Bretagne occidentale-Brest. 26, 211, 216, 221, 226, 236, 278, 311, 312, 313, 373
- DEMPSTER, A. P. (1968). A generalization of bayesian inference. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 205–247. 254, 306
- DENÈGRE, J. et SALGÉ, F. (1996). *Les systèmes d'information géographique*. Que sais-je ? Presses Universitaires de France - PUF. 25, 30, 38, 39, 41
- DESCHAMPS, E., de LUCENAY, ROSS, C., CHEVIGNY et PICARD-DOBRÉ (1966-1871). Plan général de la ville de Paris et de ses environs comprenant les Bois de Boulogne et de Vincennes. 78

- DESENDER, S. (2009). Histoatlas. <http://www.histoatlas.org>. Accessed : 2013-11-25. 53, 56
- DEVOGELE, T. (2000). Mesure d'exactitude et processus de fusion à l'aide de la distance de fréchet discrète. *Revue internationale de Géomatique*, 10(3-4):359–381. 289
- DIDIER, M. (1990). *Utilité et valeur de l'information géographique*. Economica. 39
- DONNELLY, M. et GUIZZARDI, G. (2012). *Formal Ontology in Information Systems : Proceedings of the Seventh International Conference (FOIS 2012)*, volume 239. IOS Press. 199
- DOWSLAND, K. A. et THOMPSON, J. M. (2012). Simulated annealing. *In Handbook of Natural Computing*, pages 1623–1655. Springer. 276
- DRÉO, J., PÉTROWSKI, A., TAILLARD, É. et SIARRY, P. (2003). *Métaheuristiques pour l'optimisation difficile*. Eyrolles (Editions). 264
- DUBLIN CORE METADATA INITIATIVE (2012). Dublin core metadata element set, version 1.1. Rapport technique, Dublin Core Metadata Initiative. 52, 180
- DUBOIS, D., HADJALI, A. et PRADE, H. (2003). Fuzziness and uncertainty in temporal reasoning. *J. UCS*, 9(9):1168. 199
- DUBOIS, D. et PRADE, H. (1983). Ranking fuzzy numbers in the setting of possibility theory. *Information sciences*, 30(3):183–224. 202, 206, 343, 344, 346
- DUCOM, E. (2009). Permanence des formes urbaines à travers le modèle des fringe belts. *Pérennité urbaine, ou la ville par-delà ses métamorphoses, vol.3 : Essence*, 3:51–59. 19, 32
- DURIS, D., DAWAR, A., SEGOUFIN, L., BAZGAN, C., DELON, F. et GRANDJEAN, E. (2009). *Acyclicité des hypergraphes et liens avec la logique sur les structures relationnelles finies*. Thèse de doctorat, PhD thesis, Université Paris Diderot-Paris 7. 245
- DUVERGIER, J. B. (1836). *Collection complète des lois, décrets d'intérêt général, traités internationaux, arrêtés, circulaires, instructions, etc*, volume 16. Recueil Sirey. 82
- EUZENAT, J., SHVAIKO, P. et al. (2007). *Ontology matching*, volume 333. Springer. 241
- FISHER, P., COMBER, A. et WADSWORTH, R. (2005). Nature de l'incertitude pour les données spatiales. *In DEVILLERS, R. et JEANSOULIN, R., éditeurs : Qualité de l'information géographique*, volume 1 de *Traité IGAT*, pages 43–60. Hermes. 45, 46, 48
- FISHER, P. F. (1999). Models of uncertainty in spatial data. *Geographical information systems*, 1:191–205. 45, 46

-
- FORTEMPS, P. et ROUBENS, M. (1996). Ranking and defuzzification methods based on area compensation. *Fuzzy sets and systems*, 82(3):319–330. 201, 203, 206, 343, 344
- FOURQUIER, A. (2007). *Le Paris du Premier Empire par Nicolas Maire*. Magazine du bibliophile. 70
- FOUSKAKIS, D. et DRAPER, D. (2002). Stochastic optimization : a review. *International Statistical Review*, 70(3):315–349. 265
- FRAISSE, P. (1967). *Psychologie du temps*. Presses universitaires de France. 188
- FRANK, A. U. (1998). Different types of " times " in gls. *Spatial and temporal reasoning in geographic information systems*, page 40. 188
- FREELING, A. N. (1980). Fuzzy sets and decision analysis. *Systems, Man and Cybernetics, IEEE Transactions on*, 10(7):341–354. 200
- GALINIÉ, H. (2000). *Ville, espace urbain et archéologie*. Presses universitaires François-Rabelais. 31
- GALLOIS, L. (1909). L'académie des sciences et les origines de la carte de cassini. *In Annales de Géographie*, pages 289–310. JSTOR. 118
- GALTON, A. (2004). Fields and objects in space, time, and space-time. *Spatial cognition and computation*, 4(1):39–68. 220, 231
- GAUTHIEZ, B. et ZELLER, O. (2010). Ordre textuel et ordre spatial à lyon à l'époque moderne : du parcours de visite au rôle nominal, une spatialité implicite. *Histoire et Mesure*, 1(25). 18, 44
- Gazette des Tribunaux (1830). Gazette des tribunaux, édition du 10 mars 1830. 84
- Gazette des Tribunaux (1833). Gazette des tribunaux, édition du 24 juillet 1833. 87
- Gazette des Tribunaux (1842a). Gazette des tribunaux, édition du 11 novembre 1842. 85, 87, 88, 98
- Gazette des Tribunaux (1842b). Gazette des tribunaux, édition du 7 novembre 1842. 98, 100
- Gazette des Tribunaux (1842c). Gazette des tribunaux, édition du 8 novembre 1842. 99
- GEORGE, B. et SHEKHAR, S. (2006). Time-aggregated graphs for modeling spatio-temporal networks. *In Advances in conceptual modeling-theory and practice*, pages 85–99. Springer. 314
- GILES M. FOODY, P. M. et ATKINSON (2002). *Current Status of Uncertainty Issues in Remote Sensing and GIS*, pages 287–302. Wiley Online Library. 45
- GIRARD, P.-S. et VERNIQUET, E. (1810-1819). Plan de nivellement général de la ville de Paris rapporté sur l'Atlas du plan général de la ville de Paris de Verniquet. 38
-

- GIRARD, X. (1830). Plan de la ville de Paris, divisé en 12 arrondissements et 48 quartiers... Dressé par X. Girard,..., publié en 1820, revu et considérablement augmenté en 1830. BNF, GE C-1048. 89
- GIRRES, J.-F. (2012). *Modèle d'estimation de l'imprécision des mesures géométriques de données géographiques. Application aux mesures de longueur et de surface*. Thèse de doctorat, Université Paris-Est. 115, 129, 133
- GODDING, P. (1975). *Introduction à la critique historique ; destinée aux futurs juristes*. Louvain, Cercle de Droit. 27
- GOLDBERG, D. W., WILSON, J. P. et KNOBLOCK, C. A. (2007). From text to geographic coordinates : the current state of geocoding. *URISA journal*, 19(1):33–46. 44
- GOMBOSI, M., ZALIK, B. et KRIVOGRAD, S. (2003). Comparing two sets of polygons. *International Journal of Geographical Information Science*, 17(5):431–443. 249, 252, 253
- GOMBOUST, J. (1652). Lutetia, paris, 1652 : [fac-similé, réduction de 1/4] ([reprod. en fac-sim.]). 142
- GONZALEZ-QUIJANO, L. (2012). *Filles publiques et femmes galantes : des sexualités légitimes et illégitimes à l'intérieur des espaces sociaux et géographiques parisiens (1851-1914)*. Thèse de doctorat, École des Hautes Études en Sciences Sociales. 19
- GOODCHILD, M. F. (2008). *Combining space and time : new potential for temporal GIS*, pages 179–198. Volume 1 de Knowles et Hillier (2008), 1st édition. 26
- GORDON, C. E. (2008). *Mapping decline : St. Louis and the fate of the American city*. Univ of Pennsylvania Press. 53
- GRAFF, H. F. et BARZUN, J. (1985). *The Modern Researcher*. Harcourt College Publishers. 27
- GREENWOOD, P. E. et NIKULIN, M. S. (1996). *A guide to chi-squared testing*, volume 280. John Wiley & Sons. 160
- GREGORY, I. N. et HEALEY, R. G. (2007). Historical GIS : structuring, mapping and analysing geographies of the past. *Progress in Human Geography*, 31(5):638–653. 40, 53, 58
- GREGORY, J., WEBER, K. T. et PANDA, S. (2005). Accurate mapping of ground control points for image rectification and holistic planned grazing preparation. *Methods*. 133
- GRENON, P. et SMITH, B. (2004). Snap and span : Towards dynamic spatial ontology. *Spatial cognition and computation*, 4(1):69–104. 212
- GRIBAUDI, M. (2009). Ruptures et continuités dans l'évolution de l'espace parisien. L'îlot de la Trinité entre les XVIIIe et XIXe siècles. *Histoire et Mesure*, 24:1–18. 19, 20, 285, 295, 369

-
- GROSSO, E. (2010). Integration of historical geographic data into current georeferenced frameworks : A user-centred approach. *e-Perimtron*, 5(3). 43, 142, 144
- GRUBBS, F. E. (1950). Sample criteria for testing outlying observations. *The Annals of Mathematical Statistics*, pages 27–58. 152
- GUERRA, F. (2000). 2w : new technologies for the georeferenced visualisation of historic cartography. *International Archives of Photogrammetry and Remote Sensing*, 33(B5/1 ; PART 5):339–346. 137
- GUÉVEL, S. (2004). Les canaux parisiens comme éléments fondateurs de la forme urbaine au xix^{ème} siècle? la places des marais. *Compte rendu de la séance de l'atelier "Paris", Infrastructures territoriales et urbanisation*. 92
- HAUSSMANN, G.-E. (1890-1893). *Mémoires du Baron Haussmann*. Bibliothèque Nationale de France. BNF, FRBNF34032054. 78
- HENDERSON, D., JACOBSON, S. H. et JOHNSON, A. W. (2003). The theory and practice of simulated annealing. *In Handbook of metaheuristics*, pages 287–319. Springer. 266
- HERRAULT, P.-A., FAUVEL, M., SHEEREN, D., MONTEIL, C. et PAEGELOW, M. (2013). A comparative study of geometric transformation models for the historical "map of france" registration. *Geographia Technica*, 17(1). 127, 132, 371
- HORNSBY, K. et EGENHOFER, M. J. (2000). Identity-based change : a foundation for spatio-temporal knowledge representation. *International Journal of Geographical Information Science*, 14(3):207–224. 212, 231
- HUET (1889). *Les Travaux de Paris, 1789-1889. Atlas dressé sous l'administration de M. E. Poubelle, Préfet de la Seine, sous la direction de M. A. Alphand,... par les soins de M. Huet,... M. Humblot,... M. Beckmann,... M. Fauve,... M. F. de Mallevoe,...* Impr. nationale. 37, 369
- HÄGERSTRAND, T. (1970). What about people in Regional Science? *Papers in Regional Science*, 24(1):6–21. 26, 314
- IGN (2011). Bd ortho v2. descriptif de contenu. 141
- INGBER, L. (1989). Very fast simulated re-annealing. *Mathematical and computer modelling*, 12(8):967–973. 307
- ISAKSEN, L., BARKER, E., KANSA, E. C. et BYRNE, K. (2012). Gap : a neogeo approach to classical resources. *Leonardo*, 45(1):82–83. 51, 53, 55
- ISO (2002). ISO 19108 - Information géographique - Schéma temporel. Rapport technique, International Organization for Standardization. 184
- ISO (2003). ISO 1907 - Information géographique - Schéma spatial. Rapport technique, International Organization for Standardization. 184
-

- ISO (2005). ISO 19109 - Information géographique - Règles de schéma d'application. Rapport technique, International Organization for Standardization. 184
- ISO (2014). ISO 19115 - Information géographique - Métadonnées. Rapport technique, International Organization for Standardization. 180
- JACOBET, T. (1826). Atlas général de la ville de paris, de ses faubourgs et de ses monuments; par th. jacobet. in-8 °. prospectus. bnf, notice frbnf36394280. 85
- JACOBET, T. (1827). Atlas général de la Ville de Paris/par Th. Jacobet; gravé par A. Niquet; Écrit par Hacq. BHVP, FM AT 12. 86
- JACOBET, T. (1827-1839). Atlas général de la ville, faubourgs et monuments de Paris. BHVP, ATLAS USUEL 11. 37, 38
- JACOBET, T. (1833). Atlas général de la ville, des faubourgs et des monuments de Paris. par Th. Jacobet. BNF, notice FRBNF40709634. Plan incomplet contenant des parties des trois premières livraisons de l'atlas : les feuilles 3 à 7, 10 à 12 et 16 à 18. 88
- JACOBET, T. (1842). *Affaire Hourdequin. Réponse au plaidoyer de Me Chaix d'Est-Ange (suite). Direction du bureau des plans. Organisation. 1re période 1822-23-24-25-26. 2e 1827-28. 3e 1829-30, par Th. Jacobet,...* Théodore Jacobet. 98
- JACOBET, T. (1860). Plan de Paris par Thre. Jacobet, extrait du Grand atlas de la ville de Paris en 54 feuilles. Revu et porté jusqu'aux fortifications par P. M. Dalmont, architecte, comprenant la circonscription des 20 arrondissements et celle des 80 quartiers; Gravé par Beaurain : 1860. BHVP, A 743. 105
- JAILLOT, B.-A. (1775). Nouveau plan de la ville et Faubourgs de Paris. 70, 142
- JAIN, R. (1977). A procedure for multiple-aspect decision making using fuzzy sets. *International Journal of Systems Science*, 8(1):1-7. 201, 343, 344
- JENNY, B. et HURNI, L. (2011). Studying cartographic heritage : Analysis and visualization of geometric distortions. *Computers & Graphics*, 35(2):402-411. 134, 151, 310
- JENSEN, C. S. et SNODGRASS, R. T. (1996). Semantics of time-varying information. *Information Systems*, 21(4):311-352. 188
- JIANG, B., ZHAO, S. et YIN, J. (2008). Self-organized natural roads for predicting traffic flow : a sensitivity study. *Journal of statistical mechanics : Theory and experiment*, 2008(07):P07008. 287
- JOHNSON, I. (2004). Putting time on the map : Using timemap for map animation and web delivery. *GeoInformatics*. 53, 56
- Journal Général de l'Imprimerie et de la Librairie (1825). Bibliographie de la France ou Journal Général de l'Imprimerie et de la Librairie. Accessed : 2015-11-18. 84

- JOUSSELIN, J. (1850). *Traité des servitudes d'utilité publique ou des modifications apportées par les lois et par les règlements à la propriété immobilière en faveur de l'utilité publique*. Numéro vol. 2. Videcoq,editeur. 81
- KAUPPINEN, T., MANTEGARI, G., PAAKKARINEN, P., KUITTINEN, H., HYVÖNEN, E. et BANDINI, S. (2010). Determining relevance of imprecise temporal intervals for cultural heritage information retrieval. *International journal of human-computer studies*, 68(9):549–560. 200
- KELLY, E. (2002). Digital map making at the atlas of historical county boundaries. *Origins*. 52, 53, 54
- KERRE, E. E. (1982). The use of fuzzy set theory in electrocardiological diagnostics. *Approximate reasoning in decision analysis*, pages 277–282. 201, 203, 343, 344
- KIM, K. et PARK, K. S. (1990). Ranking fuzzy numbers with index of optimism. *Fuzzy sets and Systems*, 35(2):143–150. 201, 206, 343, 344
- KIRKPATRICK, S., GELATT, C. D., VECCHI, M. P. *et al.* (1983). Optimization by simulated annealing. *science*, 220(4598):671–680. 266, 276
- KITCHIN, R., GLEESON, J. et DODGE, M. (2013). Unfolding mapping practices : a new epistemology for cartography. *Transactions of the Institute of British Geographers*, 38(3):480–496. 32
- KLEIN, E., SPIRO, M. et PATY, M. (1994). *Le temps et sa flèche*. Editions Frontières. 188
- KLIR, G. J. et YUAN, B. (1995). *Fuzzy sets and fuzzy logic*. Prentice Hall New Jersey. 45
- KNOWLES, A. K. et HILLIER, A., éditeurs (2008). *Placing history : how maps, spatial data, and GIS are changing historical scholarship /*. ESRI Press, Redlands, Calif. :, 1st édition. 5, 27, 40, 59, 355
- KOHONEN, J. (1999). A brief comparison of simulated annealing and genetic algorithm approaches. Rapport technique, Department of Computer Science University of Helsinki. 265
- KOŁODZIEJCZYK, W. (1986). Orlovsky's concept of decision-making with fuzzy preference relation-further results. *Fuzzy Sets and Systems*, 19(1):11–20. 202, 203, 206, 346
- KRAAK, M. (1999). Visualising spatial distributions. *Geographical Information Systems : Principles, Techniques, Applications and Management*. New York, John Wiley and Sons, pages 157–73. 314
- KRÄMER, S. (2010). Qu'est-ce donc qu'une trace, et quelle est sa fonction épistémologique? État des lieux. *Trivium*, 10. 23, 32
- LAFI, N. (2011). Mediterranean cosmopolitanism and its contemporary revivals. a critical approach. *New Geographies*, 5:325–335. 17, 22

- LANDRY, A. (1935). La démographie de l'ancien paris. *Journal de la société statistique de Paris*, 76:34–45. 65
- LAURENT, R., générales du ROYAUME (BELGIUM), A., des archivistes et des bibliothécaires de BELGIQUE, A. et Association des conservateurs D'ARCHIVES, d. b. e. d. m. d. B. (1979). *Sources de la géographie historique en Belgique : exposition, 26 avril - 31 mai 1979, à l'occasion du colloque organisé par les Archives générales du Royaume et l'Association des archivistes et des bibliothécaires de Belgique, 25-27 avril 1979*. Archives générales du Royaume. 28
- LAZARE, F. et LAZARE, L.-C. (1844). *Dictionnaire administratif et historique des rues de Paris et de ses monuments*. New York Public Library. 26, 89, 150, 283, 295
- LAZARE, F. et LAZARE, L.-C. (1855). *Dictionnaire administratif et historique des rues et monuments de Paris*. New York Public Library. 89, 150, 283
- LEE, E. et LI, R.-J. (1988). Comparison of fuzzy numbers based on the probability measure of fuzzy events. *Computers & Mathematics with Applications*, 15(10):887–896. 202
- LEFEBVRE, B. (2009). L'organisation parcellaire et la voirie du quartier canonial de la cathédrale de tours, du xvii^e au xviii^e siècles. *Histoire et Mesure*, 24(2). 20
- LEFEBVRE, H. (2000). *La Production de l'espace*. Ethnosociologie. Éditions Anthropos, 4^{ème} édition. 14
- LEFEVRE, E., COLOT, O. et VANNOORENBERGHE, P. (2002). Belief function combination and conflict management. *Information fusion*, 3(2):149–162. 257
- LEMAS, N. (2009). Pour une épistémologie de l'histoire urbaine française des époques modernes et contemporaines comme histoire-problème. *Histoire@Politique*, 3(9). 16
- LEPETIT, B. (1980). Histoire urbaine et espace. *Espace Géographique*, 9(1):43–54. 14
- LEPETIT, B. (1986). Espace et histoire. hommage à fernand braudel. *Annales. Économies, Sociétés, Civilisations*, 41(6):1187–1191. 15
- LEPETIT, B. (1996). La ville : cadre, objet, sujet. vingt ans de recherches françaises en histoire urbaine. *Enquête*, 4. 15, 19, 31
- LEYK, S., BOESCH, R. et WEIBEL, R. (2005). A Conceptual Framework for Uncertainty Investigation in Map-based Land Cover Change Modelling. *Transactions in GIS*, 9(3):291–322. 24, 28, 45, 47
- LI, L. et GOODCHILD, M. F. (2011a). An optimisation model for linear feature matching in geographical data conflation. *International Journal of Image and Data Fusion*, 2(4):309–328. 54

-
- LI, L. et GOODCHILD, M. F. (2011b). An optimisation model for linear feature matching in geographical data conflation. *International Journal of Image and Data Fusion*, 2(4):309–328. 251, 252, 253, 307
- LIU, T.-S. et WANG, M.-J. J. (1992). Ranking fuzzy numbers with integral value. *Fuzzy sets and systems*, 50(3):247–255. 201, 202, 203, 343, 344
- LONGLEY, P. A., GOODCHILD, M. F., MAGUIRE, D. J. et RHIND, D. W. (2005). *Geographic Information Systems and Science*. Wiley, 2 édition. 46
- LUNDY, M. et MEES, A. (1986). Convergence of an annealing algorithm. *Mathematical programming*, 34(1):111–124. 276
- MAIRE, N. (1808). *La topographie de Paris; ou Plan détaillé de la ville de Paris et de ses faubourgs*. Nicolas Maire. 72
- MANGOT, A., JACOBET, T. et BAILLY., A. (1825). Plan général de la ville et des faubourgs de Paris [Document cartographique] : Année 1825. Rapporté et dessiné par Mrs A. Mangot, attaché aux alignemens de Paris, Th. Jacobet, architecte attaché aux mêmes alignemens, et Bailly, architecte. Basé d'après le grand plan trigonométrique de feu M. Verniquet, architecte conseiller du Roi; publié par madame sa fille; Dédié et présenté à Mr le comte de Chabrol de Volvic,...; Ecrit par Martin; gravé par Adam. BHVP A454b. 84
- MANIKAS, T. W. et CAIN, J. T. (1996). Genetic algorithms vs. simulated annealing : A comparison of approaches for solving the circuit partitioning problem. Rapport technique, University of Pittsburgh. 265
- MAREUSE, E. (1925). Communication de m. edgar mareuse relative à un plan de paris au 5/1000^e qui aurait été utilisé par le baron haussmann, préfet de la seine. *Procès verbaux de la Commission du Vieux Paris*, 1:9–10. 78
- MARRAUD, M. (2010). Permanences et déplacements corporatifs dans la ville. *Histoire & Mesure*, XXV(1). 18
- MARTIN, A. (2005). La fusion d'informations. Polycopié de cours ENSIETA. 255
- MARTIN, T. et VIROL, M. (2008). *Vauban, architecte de la modernité ?* Cahiers de la Maison des sciences de l'homme Ledoux. Presses universitaires de Franche-Comté. 118
- MASCRET, A., DEVOGELE, T., LE BERRE, I. et HÉNAFF, A. (2006). *Coastline matching process based on the discrete Fréchet distance*. Springer. 289
- MASSON, M.-H. (2005). *Apports de la théorie des possibilités et des fonctions de croyance à l'analyse de données imprécises*. Thèse de doctorat, Université de Technologie de Compiègne. 192
- MATHIAN, H. (2014). *Objets géographiques et processus de changement approches spatio-temporelles*. ISTE Editions, London. 212
-

- McMASTER, R. B., LINDBERG, M. et RIPER, D. V. (2003). The national historical geographic information system (nhgis). *In Proceedings 21st International Cartographic Conference*, pages 821–828. 50
- MERLIN, P. (1827). *Répertoire universel et raisonné de jurisprudence*. Numéro vol. 10 de Répertoire universel et raisonné de jurisprudence. J. P. Roret et. 81
- MILNER, R. (2008). Bigraphs and their algebra. *Electronic Notes in Theoretical Computer Science*, 209:5–19. 226
- MONTEL, L. (2013). Les vieux quartiers de marseille au XIX^e siècle. *Histoire Urbaine*, 36(1). 19
- MORAN, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, pages 17–23. 158
- MORICONI-EBRARD, F. (1993). *L'urbanisation du monde depuis 1950*. Economica. 14
- MUSTIÈRE, S. et DEVOGELE, T. (2008). Matching networks with different levels of detail. *GeoInformatica*, 12(4):435–453. 251, 252, 253, 295, 296, 304
- MÉO, G. D. (1991). *L'Homme, la Société, l'Espace*. Anthropos. 14
- NGUYEN, D. et MEIER, A. (2006). *Introduction pratique aux bases de données relationnelles*. Collection IRIS. Springer. 36
- NOIZET, H. (2009). Les plans d'îlots vasserot, support d'un système de l'information géographique historique de paris. *EAV (Ecole nationale supérieure d'architecture de Versailles)*, 14:86–95. 73, 74, 75
- NOIZET, H., DALLO, A., BLARY, G.-X., COSTA, L., POUGET, F. *et al.* (2008). Alpage : towards the setting-up of a collaborative work tool. *Archeologia e Calcolatori*, 19:86–95. 40, 51, 52, 73, 282
- NOURANI, Y. et ANDRESEN, B. (1998). A comparison of simulated annealing cooling strategies. *Journal of Physics A : Mathematical and General*, 31(41):8373. 266, 268
- OLTEANU, A.-M. (2008). *Fusion de connaissances imparfaites pour l'appariement de données géographiques : proposition d'une approche s'appuyant sur la théorie des fonctions de croyance*. Thèse de doctorat, Université Paris-Est. 251, 252, 253, 254, 255, 258, 295, 306, 313
- PALSKY, G. (2004). Carte. 23
- PARKER, R. (2010). Historical GIS projects : Spatial data infrastructure. *Dalhousie Journal of Interdisciplinary Management*, 5. 40
- PERRET, J., GRIBAUDI, M. et BARTHELEMY, M. (2015). Roads and cities of 18th century france. *Scientific Data*, 2. Data Descriptor. 315

-
- PERROT, J.-C. (1975). *Genèse d'une ville moderne Caen au XVIIIe siècle*. Civilisations et sociétés 44. Mouton, Paris La Haye. Texte remanié de Thèse de doctorat Histoire Paris 1 1973. 15
- PEUQUET, D. J. (1994). It's About Time : A Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information Systems. *Annals of the Association of American Geographers*, 84(3):441–461. 188
- PEUQUET, D. J. (2002). *Representations of space and time*. Guilford Press. 188
- PFEIFFER, M., CARRÉ, C., DELFOSSE, V. et BILLEN, R. (2013). Virtual leodium : From an historical 3d city scale model to an archaeological information system. *In CIPA 2013 Symposium - Recording, Documentation and Cooperation for Cultural Heritage*. 65
- PINON, P. (1996). Les plans d'alignements comme source cartographique. *Bulletin du Comité Français de Cartographie*, 148:79–87. 82
- PINON, P., BOUDEC, B. L. et CARRÉ, D. (2004). *Les plans de Paris : Histoire d'une capitale*. Le Passage. 33, 35, 36, 37, 73, 82, 87
- PLEWE, B. (1997). A representation-oriented taxonomy of gradation. *In Spatial Information Theory A Theoretical Basis for GIS*, pages 121–135. Springer. 25
- PLEWE, B. (2002). The Nature of Uncertainty in Historical Geographic Information. *Transactions in GIS*, 6(4):431–456. 25, 26, 45, 46
- PLUMEJEAUD, C., COSTES, B. et GROSSO, E. (2012). Géoréférencement et appariement de données issues des cartes de cassini. *In Actes de la Conférence internationale de Géomatique et Analyse Spatiale 2012*, page 96. 137
- PONTIUS JR, R. G. et LIPPITT, C. D. (2006). Can error explain map differences over time? *Cartography and Geographic Information Science*, 33(2):159–171. 129
- PRONTEAU, J. (1986). *Edme Verniquet : 1727-1804*. Thèse de doctorat, École des Chartes. 67, 68, 69, 83, 87, 148, 170
- PRÀ, E. D. et MASTRONUNZIO, M. (2013). Rectify the river, rectify the map. geometry and geovisualization of adige river hydro-topographic historical maps. *In 8th International Workshop on Digital Approaches to Cartographic Heritage*. 137
- PUMAIN, D. (1993). La distance en géographie. *In AMAT, J.-P. et ROBIC, M.-C., éditeurs : Géographiques. Espace à vivre, espace vécu*, pages 115–134. CNDP. 14
- PUMAIN, D. et BÉGUIN, M. (2010). *La représentation des données géographiques : Statistique et cartographie*. Armand Colin. 117
- PUMAIN, D. et SAINT-JULIEN, T. (1997). *L'analyse spatiale. 1. Localisations dans l'espace*. Armand Colin. 14
- RAFFESTIN, C. (1986). *Ecogenèse territoriale et territorialité*, pages 175–185. Espaces, jeux et enjeux. Fayard et Fondation Diderot, Paris. ID : unige :4419. 21
-

- RANDELL, D. A., CUI, Z. et COHN, A. G. (1992). A spatial logic based on regions and connection. *KR*, 92:165–176. 214
- RAYWARD-SMITH, V. J., OSMAN, I. H., REEVES, C. R. et SMITH, G. D. (1996). *Modern heuristic search methods*. Wiley New York. 276
- REDMAN, C. L. et KINZIG, A. P. (2003). Resilience of past landscapes : resilience theory, society, and the longue durée. *Conservation ecology*, 7(1):14. 15
- RODIER, X. et SALIGNY, L. (2010). Modélisation des objets historiques selon la fonction, l'espace et le temps pour l'étude des dynamiques urbaines dans la longue durée. *Cybergeo*. 31
- RONCAYOLO, M. et CHESNEAU, I. (2011). *L'abécédaire de Marcel Roncayolo*. INFOLIO, infolio édition. 14, 20
- ROSSI-DORIA, O., SAMPELS, M., BIRATTARI, M., CHIARANDINI, M., DORIGO, M., GAMBARDELLA, L. M., KNOWLES, J., MANFRIN, M., MASTROLILLI, M., PAECHTER, B. *et al.* (2003). A comparison of the performance of different metaheuristics on the timetabling problem. *In Practice and Theory of Automated Timetabling IV*, pages 329–351. Springer. 265
- RUAS, A. (2010). Modèle de généralisation de données urbaines à base de contraintes et d'autonomie. *Cybergeo : European Journal of Geography*.
- RUMSEY, D. (2002). Historical maps in gis. *In* KNOWLES, A. K., éditeur : *Past Time, Past Place : GIS for History*, chapitre 1, pages 1–18. ESRI Press, 1st édition. 42, 43
- RUMSEY, D., PANORAMAS, Q., BROWSER, L., BROWSER, T. D. G., BROWSER, T. D. G., EXTRAORDINAIRE, C., CATALOGS, R. M. et CLOSE, M. U. (2002). *David Rumsey map collection*. Cartography Associates. 51, 52
- SALITCHEV, K. A. (1978). Cartographic communication/its place in the theory of science. *Cartographica : The International Journal for Geographic Information and Geovisualization*, 15(2):93–99. 32, 33, 34, 369
- SAMAL, A., SETH, S. et CUETO 1, K. (2004). A feature-based approach to conflation of geospatial sources. *International Journal of Geographical Information Science*, 18(5):459–489. 251, 252, 253, 254, 255, 290
- SCHAER, P., SKALOUD, J. et VALLET, J. (2003). Georeferencement direct des images par gps/ins-approche rigoureuse pour le calibrage des excentricités angulaires (boresight calibration). *GEOMATIK SCHWEIZ*, pages 440–445. 122
- SCHOCKAERT, S., DE COCK, M. et KERRE, E. E. (2008). Fuzzifying allen's temporal interval relations. *Fuzzy Systems, IEEE Transactions on*, 16(2):517–533. 199
- SERVICE DE GÉODÉSIE ET DE NIVELLEMENT (2009). Systèmes de projections. Rapport technique, IGN. 135

-
- SERVICE DE GÉODÉSIE ET DE NIVELLEMENT (2013). Que sont les projections lambert-93 et coniques conformes 9 zones? Rapport technique, IGN. 137
- SEVASTIANOV, P. (2007). Numerical methods for interval and fuzzy number comparison based on the probabilistic approach and dempster–shafer theory. *Information Sciences*, 177(21):4645–4661. 202
- SHAFER, G. (1976). *A mathematical theory of evidence*, volume 1. Princeton university press Princeton. 254, 311
- SHAFER, G. (1986). The combination of evidence. *International Journal of Intelligent Systems*, 1(3):155–179. 257
- SHNAYDER, E. (2010). Spatial history project. Rapport technique, Spatial History Lab. 53
- SHOEMAKER, R., HITCHCOCK, T., DAVIES, M., HOWARD, S. et McLAUGHLIN, J. (2011). Locating london’s past. <http://www.locatinglondon.org/>. Accessed : 2013-11-25. 55
- SICARD, M. (1999). Du de visu à l’in situ : la production du monument par sa représentation. *Les cahiers de médiologie*. 31
- SINTON, D. (1978). The inherent structure of information as a constraint to analysis : mapping thematic data as a case study. *Papers on Geographic Information Systems*, 6. 26
- SMETS, P. (1989). Constructing the pignistic probability function in a context of uncertainty. *In UAI*, volume 89, pages 29–40. 256
- SMETS, P. (1990). The combination of evidence in the transferable belief model. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 12(5):447–458. 256, 257, 260
- SMITH, B. (1995). On drawing lines on a map. *In Spatial Information Theory A Theoretical Basis for GIS*, pages 475–484. Springer. 25
- SMITH, B. et VARZI, A. C. (2010). Fiat and bona fide boundaries. *Philosophy and phenomenological research*, 60(2):401–420. 25
- SNODGRASS, R. T. (1992). Temporal databases. *In Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, pages 22–64. Springer. 189
- SOCIÉTÉ DE L’HISTOIRE DE PARIS ET DE L’ÎLE DE FRANCE (1881). Mémoires de la société de l’histoire de paris et de l’Île de france. 83
- SOUCHON, C. (2002). Philibert vasserot et les atlas des quartiers de paris. *Le Monde des cartes*, 171:37–41. 73, 75
- SOUTHALL, H. (2012). Rebuilding the great britain historical gis, part 3 : Integrating qualitative content for a sense of place. *Historical Methods : A Journal of Quantitative and Interdisciplinary History*. 55

- SOUTHALL, H. et AUCOTT, P. (2009). A vision of britain through time. <http://www.visionofbritain.org.uk/>. 50
- SPITZER, A. (2000). Carlo Ginzburg (1999), History, Rhetoric and Proof. *Rethinking History*, 4(2):217–220. 34
- STEFANI, C. (2010). *Maquettes numériques spatio-temporelles d'édifices patrimoniaux Modélisation de la dimension temporelle et multi-restitutions d'édifices*. Thèse de doctorat, École Nationale Supérieure d'Arts et Métiers. 31, 65, 231, 314
- STELL, J., DEL MONDO, G., THIBAUD, R. et CLARAMUNT, C. (2011). Spatio-temporal evolution as bigraph dynamics. *In Spatial Information Theory*, pages 148–167. Springer. 226
- STELL, J. G. (2003). Granularity in change over time. *Foundations of geographic information science*, pages 95–115. 213
- STOCK, O. (1998). *Spatial and temporal reasoning*. Springer. 188
- STRANO, E., NICOSIA, V., LATORA, V., PORTA, S. et BARTHÉLEMY, M. (2012). Elementary processes governing the evolution of road networks. *Scientific Reports*, 2. 20
- TALBI, E.-G. (2002). A taxonomy of hybrid metaheuristics. *Journal of heuristics*, 8(5):541–564. 265
- TANG, H.-C., HSIEH, K.-H. et CHANG, H. (2010). Experimental comparison of ranking functions in fuzzy decision making. *Journal of Information and Optimization Sciences*, 31(5):1127–1136. 201
- TARDIEU, A. (1838). Plan de Paris en 1838, avec le tracé de ses anciennes enceintes ; augmenté de tous les changements survenus jusqu'à ce jour, par Ambroise Tardieu, membre de la Commission centrale de la Société de Géographie de Paris. BNF, GED-662. 89
- TORRE, A., CALAFAT, G. et PUMA, G. (2008). Un "tournant spatial" en histoire ? : Paysages, regards, ressources. *Annales*, 63(5):1127–1144. 14
- TSENG, T. Y. et KLEIN, C. M. (1989). New algorithm for the ranking procedure in fuzzy decision-making. *Systems, Man and Cybernetics, IEEE Transactions on*, 19(5):1289–1296. 202, 203
- TUCCI, M., GIORDANO, A. et RONZA, R. W. (2010). Using spatial analysis and geovisualization to reveal urban changes : Milan, italy, 1737–2005. *Cartographica : The International Journal for Geographic Information and Geovisualization*, 45(1):47–63. 137
- VAN DAELE, K. (2014). Storing imperfect temporal information about cultural heritage in a relational database. <https://samm.univ-paris1.fr/IMG/pdf/paris2014.pdf>. 186, 281

-
- VERNIQUET, E. (1793-1799). Atlas national de la ville de Paris. Atlas du plan général de la ville de Paris levé géométriquement sur la méridienne de l'Observatoire par le Citoyen Verniquet. BHVP, FOL AT 123. 37, 38
- VILLE DE PARIS (2013). Filiation des voies actuelles. Rapport technique, open-data.paris.fr. 150
- Ville de PARIS, D. d. l. (2008). Conditions générales d'utilisation des fichiers de données cartographiques produits par la direction de l'urbanisme de la ville de paris. 140
- VOLZ, S. (2006). An iterative approach for matching multiple representations of street data. *In ISPRS Workshop, Multiple representation and interoperability of spatial data, Hanover, Germany*, pages 22–24. 251, 252, 253
- von LÜNEN, A. (2013). Tracking in a new territory : Re-imaging gis for history. *In von LÜNEN, A. et TRAVIS, C., éditeurs : History and GIS*, pages 211–239. Springer Netherlands. 24
- WALTER, F. (2005). Leitmotive, temps et contretemps de l'urbain. les lectures de villes de marcel roncayolo. *Cybergéo : European Journal of Geography*. 15, 16, 22
- WALTER, V. et FRITSCH, D. (1999a). Matching spatial data sets : a statistical approach. *International Journal of Geographical Information Science*, 13(5):445–473. 252, 253
- WALTER, V. et FRITSCH, D. (1999b). Matching spatial data sets : a statistical approach. *International Journal of Geographical Information Science*, 13(5):445–473. 251
- WANG, X. et KERRE, E. E. (2001). Reasonable properties for the ordering of fuzzy quantities (i). *Fuzzy Sets and Systems*, 118(3):375–385. 200, 201, 203
- WANG, X. et RUAN, D. (1995). On transitivity of fuzzy preference relations in ranking fuzzy numbers. *In Fuzzy Set Theory and Advanced Mathematical Applications*, pages 155–173. Springer. 200
- WANG, Y.-M. et LUO, Y. (2009). Area ranking of fuzzy numbers based on positive and negative ideal points. *Computers & Mathematics with Applications*, 58(9):1769–1779. 202
- WEISS, V. (2009). *Cens et rentes à Paris au Moyen Âge*. Histoire et archives. Hors-série. H. Champion. 18, 35
- WORBOYS, M. (1990). Reasoning about gis using temporal and dynamic logics. *In Temporal GIS Workshop. University of Maine*. 188
- WORBOYS, M. F. (1998). A generic model for spatio-bitemporal geographic information. *Spatial and temporal reasoning in geographic information systems*, pages 25–39. 189

- YAGER, R. R. (1981). A procedure for ordering fuzzy subsets of the unit interval. *Information sciences*, 24(2):143–161. 201, 202, 203, 205, 248, 343, 344
- YAGER, R. R. (1987). On the dempster-shafer framework and new combination rules. *Information sciences*, 41(2):93–137. 256
- ZADEH, L. A. (1965). Fuzzy sets. *Information and control*, 8(3):338–353. 189
- ZANOLA, S., FABRIKANT, S. I. et ÇÖLTEKIN, A. (2009). The effect of realism on the confidence in spatial data quality in stereoscopic 3d displays. *In Proceedings of the 24th international cartography conference (ICC 2009), Santiago, Chile*. 133
- ZHANG, Z. (1993). Le problème de la mise en correspondance : l'état de l'art. Rapport technique, Institut national de recherche en informatique et en automatique. 122
- Éric GROSSO (2009). Proposition pour une gestion unifiée des données anciennes. *In Colloque International de Géomatique et d'Analyse Spatiale, Spatial Analysis and GEomatics (SAGÉO 2009), Paris, France*. 181
- ØYVIND, E. (2013). Why maps are silent when texts can speak. detecting media differences through conceptual modelling. *In Proceedings of the Twenty-Sixth International Cartographic Conference*, pages 1–13, Dresden, Germany. 29

Table des figures

1.1	Premier niveau d'intégration de l'espace urbain : c'est un cadre servant de contexte, implicitement situé dans le temps par le thème de l'étude. La dimension temporelle est dominante. T	17
1.2	Second niveau d'intégration de l'espace urbain : l'espace devient un outil d'analyse des données sociales permettant de les répartir dans la ville. Les dimensions spatiales et temporelles sont considérées conjointement (S+T)	18
1.3	Transformations des structures morphologiques de l'îlot de la Trinité à Paris sur trois périodes : au milieu du XVIII ^e , en 1845 et après les percées haussmanniennes. Illustration extraite de Gribaudi (2009). . .	20
1.4	Troisième niveau d'intégration de l'espace urbain : l'espace en transformation constante est mis en rapport avec les phénomènes sociaux étudiés, chacun améliorant la compréhension de l'autre. L'espace devient dynamique, l'analyse spatio-temporelle (ST)	21
1.5	Les transformations de la rue actuelle <i>Au Maire</i> comme décrite par des plans successifs de 1716 à nos jours. L'église Saint Nicolas des Champs, au sud du Conservatoire National des Arts et Métiers, sert de repère.	26
1.6	Typologie des sources d'informations géohistoriques sur l'espace urbain ancien	28
1.7	Deux visions différentes du même espace urbain. Extraits de plans de Paris sur l'île de la cité. Les dates indiquées correspondent à la date de première publication connue.	29
1.8	Extraits de sources écrites porteuses d'informations spatiales. Source : LaDéHis, EHESS.	30
1.9	Sources iconographiques	31
1.10	Le processus de communication cartographique de Salitchev (1978)	33
1.11	Le processus de communication cartographique de Salitchev (1978) adapté pour la lecture de traces historiques dans des cartes.	34
1.12	Différentes temporalités au sein d'une même carte : les travaux de voirie effectués et édifices construits de 1789 à 1889, d'après Huet (1889)	37
1.13	Réutilisation du plan de Verniquet (à gauche) dans deux plans plus tardifs (à droite)	38
1.14	Les deux modes de représentation de l'information géographique. Raster en (a), vecteur en (b).	42

1.15	Schéma des différents filtres (U1,U2,U3) déformant la réalité perçue et représentées dans un SIG. Chaque filtre génère des imperfections différentes affectant l'information spatiale transmise.	47
1.16	Taxonomie des imperfections de l'information géohistorique, d'après De Runz (2008)	48
1.17	Extrait du SIG <i>Mapping Jewish Communities</i> (de Lange <i>et al.</i> , 2013) et légende associée. Les lieux de peuplement des communautés juives de l'empire Byzantin sont symbolisés en fonction des incertitudes associées à leurs composantes sémantiques, spatiales et temporelles. . .	53
1.18	Liens de filiations entre unités administratives dans le SIG <i>China GIS</i> Bol (2007). Issue de Berman (2013).	54
1.19	Modèle d'objets géohistoriques asynchrones et couches fusionnées dans le <i>China GIS</i>	55
1.20	Tableau récapitulatif des fonctionnalités de 15 SIG historiques existants. ■ : Non prise en compte. ■ : Prise en compte partielle ou traitement manuel. ■ : Prise en compte totale.	62
2.1	Les Buttes-Chaumont et la Butte-Aux-Cailles telles que représentées dans le plan de Maire en comparaison avec l'atlas de Verniquet et le plan de Delagrive.	71
2.2	Index des planches de l'atlas général de Jacoubet, tracé sur le plan d'assemblage, qui en constitue les feuilles 8 et 9.	86
2.3	Bornes temporelles inférieures et supérieures reportées sur l'emprise de chaque planche de l'atlas général de Jacoubet.	93
2.4	Un exemple d'incohérence entre quatre planches de l'atlas de Jacoubet, due à des périodes de dessin différentes.	94
2.5	Organisation du Bureau des Plans de Paris, depuis son déplacement du ministère de l'intérieur à la préfecture de la Seine jusqu'à sa réorganisation complète par Haussmann. Chaque ligne rouge indique une phase de réorganisation majeure, ayant des conséquences sur la production et l'utilisation des plans de Paris à la préfecture. Les couleurs indiquent la filiation du bureau des plans.	97
2.6	Filiations identifiées entre les plans de Paris utilisés dans la thèse. Certains plans non utilisés mais cités dans le chapitre sont également inclus car ils font partie de cette généalogie	108
3.1	Schéma du chapitre 3.	114
3.2	Les trois figures de la Terre, ordonnées par qualité d'approximation croissante. L'aplatissement de l'ellipsoïde est largement exagéré dans cet exemple.	116
3.3	Développement de la surface terrestre sur un cylindre. Le parallèle tangent est situé sur l'équateur (projection Mercator).	117
3.4	Le premier canevas triangulé de la France, avec la méridienne et les cinq perpendiculaires (à gauche). La position de la méridienne et de sa perpendiculaire à Paris sont visible à droite : on les voit se croiser à l'Observatoire.	119

3.5	Schéma simplifié des étapes d'une opération de triangulation géodésique.	120
3.6	Cheminement polygonal pour le relevé d'une rue. Extrait des minutes de l'atlas de Verniquet.	121
3.7	Schéma global du géoréférencement	121
3.8	Un exemple de placement de points d'amer	123
3.9	Transformations entre 1775 et 2010 du faubourg parisien du Roule. Le bouleversement total de la zone se perçoit dans l'écart entre les plans de 1808 et 1836. L'espace de 2010 n'a plus rien de commun avec celui du XVIII ^e siècle : presque aucun élément commun n'est repérable.	124
3.10	Détail du processus de géoréférencement indirect	125
3.11	Corrections géométriques effectuées lors d'un géoréférencement indirect.	126
3.12	Illustration de l'erreur résiduelle dans un cas d'estimation par moindres carrés ordinaires.	127
3.13	Illustration des effets de la répartition des points d'amer : image source et référentiel bruité	131
3.14	Illustration des effets de la répartition des points d'amer : application	131
3.15	Évolution de l'erreur résiduelle sur un échantillon de la "Carte de France", ici exprimée par l'Erreur Quadratique Moyenne résiduelle (RMSE) suivant différents modèles de transformation globale utilisés dans cette thèse et pour différents types de répartition de points d'amer. D'après les données issues de (Herrault <i>et al.</i> , 2013, p.41). . .	132
3.16	Zones Lambert I, II, III et IV. Figure issue de documents IGN.	136
3.17	Comparatif des altérations linéaires des Lambert I, II, III, IV et Lambert 93. La latitude de Paris est d'environ 48,8 degrés décimaux. Figure issue de documents IGN.	136
3.18	Référentiels géographiques envisagés.	139
3.19	Différentes orientations des plans de Paris aux XVII ^e et XVIII ^e siècles. La méridienne apparaîet en noir. Le tracé de la Seine est figuré en bleu pour aider la lecture.	143
3.20	Le carroyage de l'atlas de Verniquet (ici en version réduite au 1/8 ^{ème}) aligné sur la méridienne de Paris et sa perpendiculaire.	144
3.21	Types de points d'amer considérés. Les échelles ne sont pas constantes.	145
3.22	Schéma du processus de géoréférencement des plans anciens de Paris. Les points d'amer sont en rouge. L'erreur résiduelle des points de mesure du plan (rose) une fois géoréférencés (bleu) permet de mesurer les erreurs d'un thème cartographique.	146
3.23	Zoom sur quelques planches de l'atlas de Verniquet	148
3.24	L'observatoire de Paris, dans les minutes de l'atlas de Verniquet et dans une orthophotographie actuelle (BDORTHO®)	149
3.25	Carroyage de Verniquet dans le référentiel géographique (BDORTHO à gauche, plan parcellaire de Paris à droite)	149
3.26	Nombre de points d'amer par planches de l'atlas de Verniquet et applicabilité des transformations géométriques.	152
3.27	Moyenne et écarts-types des erreurs quadratiques moyennes sur les Planches de l'atlas de Verniquet.	153

3.28	Répartition spatiale des RMSE sur les planches de l'atlas de Verniquet dans le cas où les points d'amer sont placés sur le carroyage.	154
3.29	Vecteurs de distorsion calculés pour les points de mesures <i>bâti</i> et <i>rues</i> . Vecteurs exagérés 100 fois.	156
3.30	Schéma illustratif de la méthodologie d'analyse du plan de Verniquet	157
3.31	Types d'erreurs systématiques provoquées par une erreur de définition du carroyage dans le référentiel géographique.	158
3.32	Carte interpolée des distorsions verticales mesurées sur le bâti et les rues de l'atlas de Verniquet.	160
3.33	Corrélogramme traçant l'indice de Moran pour 19 classes de distances sur l'atlas de Verniquet, pour les distorsions verticales sur le bâti (en vert) et sur les rues (en or). Les points significatifs ($p < 0.05$) sont en rouge.	161
3.34	Carte interpolée des distorsions horizontales mesurées sur le bâti et les rues de l'atlas de Verniquet.	162
3.35	Corrélogramme traçant l'indice de Moran pour 19 classes de distances sur l'atlas de Verniquet, pour les distorsions horizontales sur le bâti (en vert) et sur les rues (en or). Les points significatifs ($p < 0.05$) sont en rouge.	163
3.36	Orientation des vecteurs de distorsions sur les bâtis et les rues par rapport à l'axe horizontal.	165
3.37	Corrélogramme traçant l'indice de Moran pour 19 classes de distances sur l'atlas de Verniquet, pour les orientations des vecteurs de distorsion sur le bâti (en vert) et sur les rues (en or). Les points significatifs ($p < 0.05$) sont en rouge.	166
3.38	Norme des vecteurs de distorsion pour les thèmes cartographiques "bâti" et "rues"	168
3.39	Orientation lissée des vecteurs de distorsion par moyenne glissante dans une fenêtre de taille v	170
3.40	Exemple de saisie du réseau viaire à l'axe de la rue sur l'atlas de Jacoubet.	172
3.41	Réseaux viaires vectorisés. Le tracé blanc indique l'état du réseau aujourd'hui.	173
4.1	Adaptation du modèle de référence de l'OGC pour les sources et observations géohistoriques.	179
4.2	Des traces aux observations géohistoriques	180
4.3	Application du schéma observation-source aux plans vectorisés.	181
4.4	Schéma des sources cartographiques.	181
4.5	Appui sur les standards de l'information géographique.	184
4.6	Représentation des traces d'entités géohistoriques par des objets cartographiques.	185
4.7	Schéma complet d'une base de données spatiale et temporelle.	186
4.8	Diagramme d'objet illustrant la représentation d'un tronçon de la rue Sainte-Opportune dans la base de données spatiale et temporelle perçue comme une observation d'une trace dans l'atlas de Verniquet.	187
4.9	Un exemple de sous-ensemble flou associé à la propriété "Jean est grand"	191

4.10	Représentation LR d'intervalles et nombres flous.	193
4.11	\widetilde{max} et \widetilde{min} de deux quantités floues.	194
4.12	Les différentes phases d'existence d'un plan topographique.	196
4.13	Le plan est temporalisé par sous-ensemble flou (a, b, α, β) , choisi pour correspondre à la période de levé topographique durant laquelle les formes urbaines tracées sur le plan sont considérées comme des représentations fiables des structures réelles. Les paramètres α et β sont des valeurs d'incertitude temporelle autour des bornes de validité du plan.	198
4.14	Quelques exemples de positions relatives entre temps flous.	200
4.15	Application des méthodes de comparaison de temps flous aux temporalités des sources parisiennes.	207
5.1	Hiérarchie des relations considérées dans le modèle de graphe spatio-temporel.	216
5.2	Situations interdites dans le graphe spatio-temporel. Figure d'après Del Mondo (2011).	216
5.3	Un exemple de graphe spatio-temporel représentant les transformations de parcelles cadastrales à trois instants. Toutes les relations spatio-temporelles ne sont pas représentées.	217
5.4	Exemple d'instanciation du modèle de graphe spatio-temporel (à gauche) en un schéma de base de données relationnel (à droite).	220
5.5	Contrainte sémantique décrivant une fusion.	220
5.6	Sens rétro-temporel : deux ordre possibles	225
5.7	Exemple de sources décrivant des observations à différents niveaux de détail.	226
5.8	Graphe géohistorique invalide résultant d'une différence de niveau de détail.	227
5.9	Amalgamation dans le cas d'une continuation multiple.	227
5.10	Situation interdite par le graphe spatio-temporel, autorisée dans le graphe géohistorique.	228
5.11	Un exemple de graphe géohistorique.	229
5.12	Un exemple de graphe géohistorique cyclique.	230
5.13	Une transformation multisource.	230
5.14	Typologie des processus spatio-temporels représentés dans le graphe géohistorique	231
5.15	Processus spatio-temporel de stabilité.	231
5.16	Processus spatio-temporel de transformation.	232
5.17	Processus spatio-temporel d'apparition.	232
5.18	Processus spatio-temporel de disparition.	232
5.19	Processus spatio-temporel de fusion.	233
5.20	Processus spatio-temporel de scission.	233
5.21	Processus spatio-temporel d'annexion.	233
5.22	Processus spatio-temporel de séparation.	234
5.23	Processus spatio-temporel de réallocation.	234

6.1	Schéma général décrivant l’instanciation d’un graphe géohistorique à partir d’un ensemble d’observations.	240
6.2	Un exemple d’observation agrégée.	244
6.3	Un hypergraphe de filiation représentant les processus spatio-temporels d’un graphe de filiation.	246
6.4	Exemples d’hypergraphes de filiation non admissibles.	246
6.5	Exemple de paramétrisation de la mesure de succession temporelle <i>Crank</i>	262
6.6	Schéma de l’algorithme de recuit simulé pour la découverte des relations de filiation.	267
6.7	Deux cas d’hyperarcs voisins	269
6.8	Illustration des modifications d’extension, de réduction, de fusion et de scission.	270
6.9	Schéma général de la création d’une modification d’un hypergraphe de filiation.	270
6.10	Exemples de l’étape de filtrage pour une modification de type ”extension”. Seul l’hyperarc concerné par la modification est montré. . .	271
6.11	Exemples de l’étape de filtrage pour une modification de type ”réduction”.	272
6.12	Exemples de l’étape de filtrage pour une modification de type ”fusion”.	273
6.13	Exemples de la scission d’un hyperarc en utilisant une coupe aléatoire dans les sous-graphes spatiaux induits par sa tête et sa queue.	274
6.14	Génération et sélection d’une solution du voisinage de la solution courante.	275
6.15	Schéma général de la phase de typage des filiations.	278
6.16	Décomposition d’un hypergraphe de filiation en graphe simple.	278
6.17	Schéma général du prototype implémenté.	282
6.18	Temps valides des différents atlas.	283
6.19	Modélisation des réseaux viaires de Paris dans la base de données spatiale et temporelle	284
6.20	Illustration des décalages planimétriques entre réseaux viaires sur l’île saint Louis et l’îlot de la trinité.	285
6.21	Les états successifs de l’îlot de la trinité et les réseaux vectorisés correspondants.	286
6.22	Les états successifs du quartier des Arts et Métiers et les réseaux vectorisés correspondants.	287
6.23	Agrégation des tronçons de rues par création de routes naturelles. Chaque couleur correspond à une polyligne.	288
6.24	Exemple de calcul de la distance de Fréchet entre deux lignes polygonales.	289
6.25	Appariement résultant de la découverte des relations de filiations entre tronçons de l’îlot de la trinité.	294
6.26	Découverte des relations de filiation avec pénalité des hyperarcs de cardinalité $n : m$	295
6.27	Graphe G non étiqueté créé par la phase de découverte des relations de filiation pour l’îlot de la Trinité.	297

6.28	Graphe de filiation G_f après typage des relations de filiations découvertes pour l'îlot de la trinité.	298
6.29	Graphe de filiation G_f après typage des relations de filiations découvertes pour l'îlot de la trinité, sans considérer le nom des rues.	299
6.30	Appariement résultant de la découverte des relations de filiations entre tronçons pour le quartier des Arts et Métiers.	301
6.31	Cas d'échecs de l'algorithme de découverte des relations de filiation.	302
6.32	Détail des situations d'échec de la phase de découverte des relations de filiation pour le quartier des Arts et Métiers.	303
6.33	Faux positifs lors de la découverte des filiations, dus au décalage planimétrique des données. À gauche, les décalages planimétriques de l'atlas de Jacoubet sont réduits manuellement par rapport à celui de Vasserot.	304
A.1	Carte des points de triangulation de l'atlas national de Verniquet.	320
B.1	<i>Atlas National de la Ville de Paris</i> par Edmé Verniquet, fac-similé en 72 planches de Marcel Legoux daté de 1983. Assemblage après géoréférencement.	324
C.1	<i>La topographie de Paris ou plan détaillé de la ville de Paris et de ses faubourgs</i> par Nicolas Maire, édition de 1808 en 20 feuilles, partie de l'atlas du même nom. Assemblage après géoréférencement.	326
D.1	<i>Atlas Général de la Ville et des Faubourgs de Paris</i> par Théodore Jacoubet, édition de 1836 en 54 planches. Les planches 1 et 2 (titre), 8 et 9 (plan d'assemblage), 46-47 et 53-54 (nomenclatures et statistiques), non cartographiques, ne sont pas affichées. Assemblage après géoréférencement.	328
E.1	Atlas de Paris en 16 feuilles, édition de 1888, réalisé par les géomètres du Service du Plan de Paris, assemblé après géoréférencement des feuilles.	330
F.1	Réseau viaire vectorisé de l'atlas de Verniquet superposé au réseau de Paris actuel issu de la BDTOPO.	332
F.2	Réseau viaire vectorisé de l'atlas de Vasserot superposé au réseau de Paris actuel issu de la BDTOPO.	333
F.3	Réseau viaire vectorisé de l'atlas de Jacoubet superposé au réseau de Paris actuel issu de la BDTOPO.	334
F.4	Réseau viaire vectorisé de l'atlas municipal superposé au réseau de Paris actuel issu de la BDTOPO.	335
H.1	Configurations de quantités floues évaluées. Exemples d'après (Bortolan et Degani, 1985) et (Chen et Lu, 2001b).	342
I.1	Configurations de temporalités floues sur lesquelles sont évaluées les méthodes calculant un degré d'antécédence.	345

Liste des tableaux

1.1	Résumé des trois niveaux de relations entre espace urbain et temps et des propriétés particulières de chaque niveau.	22
2.1	Types d'événements affectant Paris ayant permis de dater les planches de l'atlas de Jacoubet, triés par ordre décroissant de fiabilité.	89
2.2	Datation des planches de l'atlas de Jacoubet commencé en 1825 et publié en 1836, rassemblées par livraison.	90
4.1	Temporalités valides des plans utilisés dans la thèse.	199
6.1	Modélisation des processus spatio-temporels sous la forme d'hyperarcs	245
6.2	Tableau résumant l'adéquation des principales approches d'appariement de données géographiques aux fonctionnalités nécessaires pour la découverte de relations de filiations.	252
6.3	Tableau résumant l'adéquation des principales approches d'appariement de données géographiques aux fonctionnalités nécessaires pour la découverte de relations de filiation.	253
6.4	Qualité de l'appariement sur l'îlot de la trinité	296
6.5	Qualité de l'appariement sur le quartier Saint Martin	305
H.1	Évaluation des méthodes de tri d'ensembles flous sur les exemples en figure H.1. Les lignes en jaune aboutissent au même rangement que celui de référence quelque soit la situation. Les cases en gris identifient le premier cas d'échec du tri par rapport à l'ordre de référence. Partie 1.	343
H.3	Évaluation des méthodes de tri d'ensembles flous sur les exemples en figure H.1. Les lignes en jaune aboutissent au même rangement que celui de référence quelque soit la situation. Les cases en gris identifient le premier cas d'échec du tri par rapport à l'ordre de référence. Partie 2.	344
I.1	Degré d'antécédence des temporalité floues représentées dans la figure I.1 avec différentes méthodes fondées sur la dominance entre sous-ensembles flous	346

Un système d'information géographique pour le suivi d'objets historiques urbains à travers l'espace et le temps

Résumé

Les systèmes d'information géographiques (SIG) sont aujourd'hui fréquemment utilisés en sciences sociales et notamment en histoire pour leurs capacités de visualisation, de partage et de stockage de données historiques spatialisées. Ils ouvrent ainsi aux chercheurs en histoire des perspectives de croisement d'informations et d'analyse à grande échelle, non seulement de l'évolution de phénomènes spatialisés, mais aussi des interactions entre ces phénomènes et les transformations de l'espace. La représentation d'un espace en transformation au sein d'un SIG implique notamment la création et la manipulation de données spatio-temporelles. Créées à partir des sources d'informations sur l'espace ancien disponibles, celles-ci sont généralement hétérogènes et imparfaites. Les SIG existants ne permettent pas de modéliser et d'intégrer aisément des données hétérogènes et imparfaites au sein de bases de données spatio-temporelles.

Cette intégration est l'objectif de ce travail de thèse. Nous proposons une approche d'intégration de données géographiques hétérogènes et imparfaites sur l'espace urbain ancien au sein d'un modèle de données spatio-temporelles permettant l'analyse des transformations de cet espace. Ces données sont saisies à partir de plans topographiques. Au travers de plusieurs étapes, allant de l'analyse des sources cartographiques utilisées à un processus semi-automatique d'intégration de données spatio-temporelles, nous proposons une approche générique permettant d'instancier une base de données spatio-temporelles à partir de cartes historiques. Celle-ci est finalement appliquée à la création d'une base de données sur le réseau des rues de Paris des XVIII^e et XIX^e siècles à partir de plans topographiques de la capitale.

Mots-clés : Données géo-historiques urbaines, appariement spatio-temporel, intégration de données géohistoriques.



A Geographic Information System for tracking urban historical objects through space and time

Abstract

Geographic information systems (GIS) are increasingly used for leading historical studies because of their ability to display, store and share geo-historical data. They provide an opportunity for exploring and analyzing spatialized phenomena and the interactions between such phenomena and spatial dynamics. To achieve this goal, GIS have to manage spatio-temporal data describing the transformations of geographical entities. These data are also highly imperfect since knowledge about the past is only available through imprecise or uncertain historical sources such as maps. To date, no GIS is able to integrate, manage and analyze such imperfect data. In this thesis, we focus on the integration of spatio-temporal data about urban space extracted from historical topographic maps on the city of Paris. We propose a process that allows to create spatio-temporal graphs from geohistorical vector data extracted from georeferenced maps of the city. After the analysis of the maps and the measure of their spatial and temporal imperfections, we propose a spatio-temporal model named geohistorical graph and a semi-automatic spatio-temporal data matching process able to build such graphs from vector data extracted from old topographic maps. Our method is tested and validated on the street networks of Paris extracted from maps covering the period from the late XVIIIth century to the late XIXth century.

Keywords : Geo-Historical data, Spatio-temporal data matching, Geo-Historical data integration.